Comments Submitted for the Peer Review Workshop on the EPA Draft Toxicological Review Document on Perchlorate

Table of Contents

Review of Neurobehavioral Developmental Study

Charles H. Emerson, M.D., 2/1/99
Professor of Medicine
University of Massachusetts Medical School
Worcester, Massachusetts 01655

Submitted for the Peer Review workshop on the EPA Draft Toxicological Review Document on Perchlorate – 2/10/99, 8:30 AM at San Bernardino City Council Chambers, 300 North D Street, San Bernardino, California

"A Neurobehavioral Developmental Study of Ammonium Perchlorate Administered Orally in Drinking Water to Rats"

SUMMARY OF STUDY

The goal of the study was to provide information regarding neurobehavioral effects of ammonium perchlorate exposure in pregnant and recently pregnant rats and in their offspring, both during the neonatal period and after weaning.

125 female rats were acclimatized to the testing facility and divided into 5 equal groups (25 rats per group). All rats were then permitted to cohabit with males for a maximum of 7 days. After cohabitation treatment with ammonium perchlorate in the drinking water was started and continued until 10 days after they delivered or for 31 days after the first day of gestation. The target doses of ammonium perchlorate were 0 (carrier), 0.1, 1.0, 3.0, and 10 mg/kg/day. Rats were observed during ammonium perchlorate treatment for abnormal neuromotor function. At birth maternal behavior, litter sizes and pup viability was noted.

Ammonium perchlorate treated dams (F0 generation) that delivered liters of more than 8 pups per liter were observed with respect to their maternal behavior and other parameters for 22 days after delivery and then sacrificed. This group of Dams were sacrificed, therefore, about 12 days after ammonium perchlorate exposure. Dams with no surviving pups or less than 8 pups per liter were sacrificed 10 days after delivery, a time at which they would have been exposed to ammonium perchlorate.

Pups (F1 generation) were observed and evaluated after their birth. On the fifth neonatal day some pups were sacrificed and some were committed to continued observation which extended to after weaning and likely extended into a period in which there was no ammonium perchlorate exposure. These pups were assigned, using a table of random units, to one of four studies. One male and one female pup from each litter was assigned to each study.These studies were (Subset 1) Brain weights and neurohistology examination, (Subset 2) Passive avoidance and watermaze testing, (Subset 3) Motor activity and auditory startle habituation, and (Subset 4) Regional brain weight and neurohistology examinations.

Results of serum thyroid function tests (T4, T3, and TSH) were reported for dams sacrificed at day 10. The results of thyroid histopathology examinations were reported for dams sacrificed at day 10 and for pups sacrificed on neonatal day 5. Necropsy data were reported on all rats.

No deaths, clinical observations, observations of autonomic dysfunction; or necropsy observations relating to ammonium perchlorate administration during pregnancy or after delivery were noted. No clinical observations relating to test substance were noted in neonatal pups. Pups followed through a post weaning period did not have ammonium perchlorate related neuropathological changes, variances in motor activity or effects relating to the auditory startle response.

Serum TSH, T3, and T4 concentrations were higher in some or all of the ammonium perchlorate treated dams as compared to the control group and many of these changes were statistically significant but a dose-response relationship was not always observed.

There was no treatment-related microscopic changes in the thyroids of dams. Hypertrophy/hyperplasia of follicular epithelium and follicle size was altered in newborn (5 day) male pups that received 10 mg/kg/day ammonium perchlorate but not in female pups. Pups sacrificed after this time did not have ammonium perchlorate related effects on thyroid weight.

CRITIQUE

The stated goal of this study was "to provide information for use in evaluating the potential for neurobehavioral effects in offspring after exposure to ammonium perchlorate in utero and during the neonatal period". In general a group of specific questions that the study was intended to address is not provided in the main reports. Rather the specific questions must be inferred from the endpoints. Maternal hypothyroidism, fetal hypothyroidism, and particularly combined maternal-fetal hypothyroidism has major impacts on development. More discussion of this in the context of the study design and goals would have been of interest.

There are several areas in which the design of this study might be improved. More information is also needed regarding the methodology for an important endpoint relating to thyroid function.

(1)The results of the thyroid function studies are absolutely dependent upon the comparison between the control and treated groups. However, the number of rats in the control groups is less than in the treated groups. The study would have been strengthened if there were control groups for every treatment group or control (carrier) groups for at lest two of the four treatment groups instead of there being only one control group. This would help determine if the apparent effect of ammonium perchlorate treatment on serum thyroid function parameters, which sometimes was not dose-dependent, is likely to be valid.

(2) Perchorate exerts its antithyroid effect by interfering with the uptake of iodide by the thyroid. The antithyroid effects of perchlorate are probably influenced by the iodine intake of the organism. Iodine balance is more likely to be perturbed during pregnancy and is of critical importance for development. Therefore, this study would be enhanced if it provided information regarding the iodine intake of the rats.

(3) Serum TSH concentrations are widely acknowledged to be the most sensitive index of primary thyroid dysfunction. It would be helpful to have more information concerning the serum TSH assay used for this study within the main body of the report. It is important that the TSH assay employed be valid for rats. There are major differences among the TSH assays for different species. The TSH assay method cited in Appendix M is not widely used in the literature.

(4) Collection of blood in 5 day old pups for serum thyroid function tests is discussed in the study report but test results do not appear to be available. It would have been of interest to have serum TSH results in pups. This may have required changing the protocol to permit pooling of blood from pups in the same litter so that sufficient material was available for analysis.

(5) Data on the excretion of ammonium perchlorate in the milk of dams would have been of interest and relevant to the question of whether neonatal pups would have been exposed to the antithyroid effects of perchlorate.

Some areas in the study were difficult to follow or sharp enough distinctions do not appear to have been made. A distinction should be drawn with respect to the thyroid studies in 5 day old pups and pups studied during the post weaning period because the exposure to perchlorate of the two groups is likely to differ.

In general the study was conducted in accordance with good laboratory practices and the changes in the protocol during the course of the study were minor and did not impact the findings.

The significance of the findings that serum TSH, T3, and T4 concentrations were higher in some or all of the ammonium perchlorate treated dams as compared to the control group is not clear and as noted may need further study to establish its validity. However, an increase in serum thyroid hormone concentrations in association with an increase in serum TSH is not consistent with the reported anti-thyroid effects of perchlorate and of anti-thyroid drugs in general. Rather the effect of anti-thyroid drugs is to lower thyroid hormone secretion and this causes a secondary increase in serum TSH.

The study is useful for hazard characterization purposes. This is true because of its wide ranging data collection dealing with organ systems other than the thyroid and, as far at the thyroid is concerned, because thyroid weight and histology are important and valid

indices of anti-thyroid drug effects. The post weaning studies dealing with neurological development are useful for hazard characterization purposes in the context of the potential effects of ammonium perchlorate.

Review of 90-Day Subchronic Oral Bioassay Study
Charles H. Emerson, M.D., 2/1/99
Professor of Medicine
University of Massachusetts Medical School
Worcester, Massachusetts 01655

"90-Day Subchronic Oral Bioassay Study"

SUMMARY OF STUDY

This study was performed male and female rats. Its goal was to the "evaluate the potential toxicity of ammonium perchlorate" when administered in drinking water in concentrations designed to deliver a range of doses from 0.01 to 10 mg/kg/day. Groups of rats, consisting of 10 rats per group, received 0, 0.01, 0.05, 0.2, 1.0, and 10 mg/kg/day. For various groups the ammonium perchlorate treatment duration was 14 days, 90 days, or 90 days followed by a recovery period of 30 days. The effect of recovering from ammonium perchlorate administration was tested only in rats that received no ammonium perchlorate (control group), and rats treated to receive target dose levels of 0.05, 1.0, and 10.0 mg/kg/day. Rats were observed daily for signs of toxicity and their weight and water consumption was measured weekly. Euthanasia was performed after 14 and 90 days of ammonium perchlorate treatment or after 30 days of recovery from the 90 day ammonium perchlorate treatment. Vaginal smears were examined three weeks prior to euthanasia to evaluate the estrous cycle in female rats treated for 90 days and in those after the 30 day recovery period. In male rats treated for a similar duration, semen samples were obtained for sperm count, sperm motility, and sperm morphology. In all rats eye examinations were performed at the beginning to the study, near the time that ammonium perchlorate administration was being completed, and near the end of recovery. Bone marrow samples were collected from all animals that received euthanasia at 90 and 120 days. The samples were evaluated for micronuclei formation. Blood samples were obtained at the time of euthanasia for CBC, routine blood chemistries, and thyroid function tests consisting of serum TSH, serum T3 and serum T4. A complete gross necropsy was performed on all rats, organ weights were recorded and tissues were preserve for pathologic examination.

No clinical signs of toxicity were noted in this study. One female rat that received 0.05 mg/kg/day per kg per day of ammonium perchlorate (the second lowest dose employed) was found dead during recovery seven days after ammonium perchlorate was stopped. There was no meaningful difference in food or water consumption amongst the various groups of rats. The intake of perchlorate was similar to the intended dose, and weight gain was similar in all groups with the exception of the group of male rats that received 10 mg/kg/day per kg per day of ammonium perchlorate. Statistically significant differences in the CBC and blood chemistries, testes weight and organ weights were noted in a few groups but these differences did not have any relationship to the ammonium perchlorate dose. There was no difference between control and ammonium perchlorate treated rats with respect to the eye examination, the estrous cycle or the sperm count. The necropsy observations were unremarkable with the exception of mild

reddening of the thyroid gland in three males that received 10 mg/kg/day kg per mg per day of ammonia perchlorate. No evidence of a mutagenic effect, as assessed by the bone marrow examination for micronuclei formation, was noted.

In the male and female rats that received euthanasia at 14 days, serum TSH concentrations were significantly higher in rats receiving 0.2, 1.0, 10.0 mg/kg/day ammonium perchlorate than in the control groups. The differences in serum TSH were related to the administered dose of ammonium perchlorate. Serum TSH concentrations were approximately 20 percent higher than that of the control groups in rats receiving 0.2 mg/kg/day ammonium perchlorate and approximately 60 percent higher than in control rats in those receiving 10 mg/kg/day ammonium perchlorate. In male and female rats that received euthanasia at 90 days, serum TSH concentrations were also higher in groups treated with 10 mg/kg/day ammonium perchlorate than in control groups. However, the differences between the control and ammonium perchlorate treated rats were only about 20 percent, differences considerably less than the 60 percent difference noted in rats treated with ammonium perchlorate for 14 days. Serum T4 and serum T3 concentrations were lower in male and female rats that received 10 mg/kg/day ammonium perchlorate for 14 or 90 days as compared to control rats. In contrast to the results seen with serum TSH, however, there was not a good relationship between the dose of ammonium perchlorate administered and the differences in serum T4 and T3 between control and treated rats.

CRITIQUE

The goal of this study as presented in the amended final report was to evaluate the potential toxicity of ammonium perchlorate when administered chronically by the oral route. In appendix "I" the stated goal is to determine the NOEL for thyroid effects by perchlorate, and to look for other potential target tissues. This latter goal more specifically states the information being sought by this study. Otherwise a group of specific questions that the study was intended to address is not provided in the main reports. However it is not difficult to ascertain many specific goals of the study as it contains numerous well identified end points. These are broad ranging, dealing with every major organ and most major Endocrine systems.

There are several areas in which the design of this study might be improved. More information is also needed regarding the methodology for an important endpoint relating to thyroid function.

(1)The results of the thyroid function studies are absolutely dependent upon the comparison between the control and treated groups. However, the number of rats in the control groups is less than in the treated groups. The study would have been strengthened if there were control groups for every treatment group or control groups for at lest three of the five treatment groups (per sex) instead of there being only one control group for each sex. This would help determine if the apparent effect of ammonium perchlorate treatment on serum T4 and T3 concentrations, which of note were not dose-dependent, is likely to be valid.

(2) Perchorate exerts its antithyroid effect by interfering with the uptake of iodide by the thyroid. The antithyroid effects of perchlorate are probably influenced by the iodine intake of the organism. Therefore, this study would be enhanced if it provided information regarding the iodine intake of the rats.

(3) Serum TSH concentrations are widely acknowledged to be the most sensitive index of primary thyroid dysfunction. It would be helpful to have more information concerning the serum TSH assay used for this study within the main body of the report. It is important that the TSH assay employed be valid for rats. There are major differences among the TSH assays for different species. The TSH assay method cited in Appendix I is not widely used in the literature.

In general the study was conducted in accordance with good laboratory practices and the changes in the protocol during the course of the study were minor and did not impact the findings. The presentation of the investigations is detailed in terms of the raw data but could be improved by providing analytic information. For example, it would be informative to express the results of serum T4, T3 and TSH for ammonium perchlorate treated groups as a percentage of the results obtained for the control groups. The rationale for the design of the study is sparsely presented and it is not clear how the chosen rationale relates to the study design. Papers by Caldwell et al. and Brabant et al. were cited (Amended Final Report, p 32) as being the basis for the dose level selection. The study by Caldwell et al. is not widely available in traditional databases such as MEDLINE or TOXLINE. The study by Brabant et al. was performed in five healthy adult men, ages 25 to 28. They were treated with 0.2 mg iodine daily for four weeks and then with 900 mg of perchlorate daily for four weeks. Thyroid function parameters were assessed at the end of the iodine treatment period and at the end of the perchlorate treatment period. The initial dose of iodine is within the range of the usual daily iodine intake in the United States. In iodine sufficient areas this should not have a major impact on thyroid function. The perchlorate dose was about 13 mg/kg/day assuming the weight of the participants was about 70 kg. This dose is somewhat higher than the highest dose administered to rats in the present study. Interestingly, despite that fact that perchlorate treatment decreased intra-thyroid iodine, serum TSH concentrations were actually lower after perchlorate treatment than values just prior to treatment and no effect of perchlorate on thyroid weight was reported by Brabant et al.

Despite the questions raised regarding serum T4, T3 and TSH determinations the study is useful for hazard characterization purposes. This is true because of its wide ranging data collection dealing with organ systems other than the thyroid and, as far at the thyroid is concerned, because thyroid weight and histology are important and valid indices of anti-thyroid drug effects. However, the dose-response relationships pertaining to the potential effects of perchlorate on the thyroid may differ between inmature and adult rats. For example, the mechanism for autoregulation of intrathyroidal iodine may not be fully developed in inmature rats or humans and, in adults, this mechanism may protect against perchlorate toxicity. Furthermore, individuals with untreated subclinical hypothyroidism

may be more sensitive to the antithyroid effects of perchlorate than persons with normal thyroid function. This may also be the case for strains of rats that are prone to developing autoimmune thyroiditis. Finally it is difficult to assess the degree to which the present study is valid for humans. As noted, perchlorate administration in doses similar to the highest dose administered in this study was not associated with an increase in serum TSH nor was there an effect on thyroid weight.

Review of Neurobehavioral Study

Prepared by R. Thomas Zoeller, Ph.D.
University of Massachusetts
Department of Biology
Amherst, MA 01003

Review of Neurobehavioral Study

1. Design The stated objective of this experiment was to provide information for use in evaluating the potential for neurobehavioral effects in offspring after exposure to ammonium perchlorate *in utero* and during the neonatal period. To accomplish this, pregnant rats were exposed to 5 doses of ammonium perchlorate (AP) dissolved in the drinking water (0, 0.1, 1, 3, 10 mg/kg/day). Dosing began on gestational day 0 (GD0, the day sperm was identified in vaginal contents) and water consumption was monitored to make adjustments in dosing necessary to achieve the target dose. AP dosing was continued throughout pregnancy to the 10th day of lactation. Litters were culled to 8 on DP5. On DP5, dams and litters were chosen for blood collection to measure thyroid hormones (T4, T3, TSH). F1 animals were used in 4 subsets as follows:

Subset 1:	DP12 - sacrifice for brain weights, and neurohistology
Subset 2:	DP23-25 & 30-32 - Passive Avoidance
	DP59-63 & 66-70 - Water maze
	DP90-92 - sacrifice with blood collection
Subset 3:	DP14, 18, 22 & 59: Motor Activity
	DP23 & 60: Startle habituation
	DP67-69 - Sacrifice with blood collection
Subset 4:	DP80-86: Sacrifice for brain weights and neurohistology

Critique:

This study clearly reflects a great deal of difficult and meticulous work. The strengths of the specific design include the number of animals used, the use of litters as the statistical "n" to avoid potential litter effects, the dose-range of perchlorate evaluated, and the measurements of water and food consumption. In addition, lab records were thorough and well-organized. However, the are a number of weaknesses in the design considering that the focus of the experiment was to evaluate the potential for neurobehavioral effects of ammonium perchlorate acting on the thyroid system. These weaknesses are listed below.

Behavioral End-Points

Most of the behavioral end-points used to evaluate perchlorate neurotoxicity are not well-characterized to be related to thyroid hormone action. For example, there are few reports of the effect of perinatal hypothyroidism on passive avoidance, but those studies available report a lack of effect of thyroid hormone on passive

avoidance . Therefore, one would not propose that perchlorate would affect passive avoidance, and no effect was observed. Likewise, the effect of perinatal hypothyroidism on the behavior of animals in the auditory habituation test does not appear to be well-characterized. The effect of thyroid dysfunction on motor activity is better characterized and appears to be a valuable end-point for this study. However, it would strengthen the ability to interpret the present data if the overt effects of hypothyroidism (produced by a thionamide goitrogen for example) had been evaluated to validate the testing procedures.

Neuroanatomical End-Points

Brain weight, or the weight of various brain regions, also are poor indicators of the potential thyroid-disrupting effects of perchlorate. Although a few studies demonstrate reduced brain or cerebellar weight during severe and persistant hypothyroidism, not all studies report this effect , suggesting that brain weight may not be a very sensitive indicator of thyroid disruption.

In contrast, myelination is well-known to be affected by thyroid hormone in both animals and humans. Thyroid hormone directly regulates the expression of the gene encoding myelin basic proteins, and the size of the corpus callosum, as well as other major myelinated structures within the central nervous system, is affected by thyroid hormone in humans and animals. Therefore, the size of the corpus callosum is a reasonable end-point for this study. However, only the caudal portion of the corpus callosum is affected by thyroid hormone and it is not clear whether the single linear measurement taken at the mid-point of the corpus callosum in this study would reflect changes in thyroid hormone action.

Thyroid Histopathology

The methods used for analysis of thyroid histopathology may not detect small changes that occur at lower doses of perchlorate. The diameter of 5 thyroid follicles was measured using an ocular micrometer on each of three 6 m sections through a single thyroid gland. These values were averaged for each thyroid and used in the statistical analysis. However, there is no description about how the 5 follicles were chosen to measure. Certainly, there are hundreds of follicle profiles within a single 6 m section. These profiles will vary in diameter both because follicles differ in size and because different amounts of the follicle are represented in a single section. Moreover, the follicles are not perfectly circular. Thus, several subjective elements may have a large impact on the results of these studies. For example, how were follicles chosen for measurement? Was the diameter the greatest distance across the colloid? Did the 3 sections taken for measurements have the same follicles on them? (This is important because it will determine whether the 15 measurements were independent). Measurements of follicular area by standard computer-assisted morphometry would appear to be a much more accurate way to obtain data indicative of changes in thyroid function. It is also likely that more measurements should be included (20 profiles/section x 5 sections insuring that the sections chosen were far enough apart that duplicate

measures of the same follicle would be avoided). In addition, performing the statistical analysis only on the average linear measurement of follicle diameter is not likely to be the most sensitive index of thyroid function. The reason is that there is considerable variation in follicle profile size within a single section through the thyroid. Follicles are spherical, a single 6 m section will intersect different follicles at different distances from their geometric center, creating profiles of different diameters and areas. This non-systematic variation could be separated from variation due to treatment effects using a frequency histogram approach in addition to the ANOVA. The Kolmogerov-Smirnov test would provide such a tool.

Perchlorate Dosing

The duration of perchlorate administration is not well-justified and is not optimal for identifying effects due to thyroid disruption. The colloid represents stored iodinated thyroglobulin, the precursor form of thyroid hormones. Therefore, it is obvious that there are large quantities of thyroid hormone precursor stored in the thyroid gland. Initiating perchlorate on the day of mating is too late to produce hypothyroidism in the dams until late in gestation, even if the dose of perchlorate used fully suppresses thyroid hormone synthesis. There are probably two reasons for this. First, as mentioned above, large amounts of thyroid hormne precursor are stored in the colloid. Second, many tissues, especially fetal brain, activate a number of compensatory responses to hypothyroxinemia that is protective against mild hypothyroxinemia. Therefore, the degree of hypothyroidism of the fetuses and pups is not likely to be representative of human exposure patterns.

It is also not clear why ammonium perchlorate exposure was terminated on lactational day 10. It is highly likely that perchlorate is selectively taken up into the milk. Mammary tissue in both rodents and humans expresses the sodium/iodide symporter which selectively transports perchlorate with an affinity that is greater than for iodide itself. In addition, iodide uptake (and thus, perchlorate uptake) is enhanced by prolactin indicating that mammary tissues concentrate iodide into milk. It is known that perchlorate is a potent inhibitor of iodide uptake into milk by a mechanism that is independent of chloride uptake ; thus, perchlorate will not follow a path into the neonate that is equivalent to other anions such as chloride. Although this reviewer could not find reports of perchlorate measurements in milk, it is known that iodide is concentrated in milk and that iodine levels in breast milk may be quite high. In addition, the gastric and intestinal mucosa of the pup is rich in sodium/iodide symporter, all suggesting that the neonate may be exposed to a higher concentration of perchlorate than the dam. It seems important to test this by dosing lactating dams throughout the period of lactation. In addition, lactating women in perchlorate-contaminated regions are likely to be contaminated throughout the period of lactation.

Recommendations on Study Design

Considering that the potential neurotoxicological effects of perchlorate are believed to be mediated solely by actions on iodide uptake into the thyroid gland resulting in reduced thyroid hormone synthesis, it is important to choose end-points known to be sensitive to thyroid hormone regulation. Because thyroid hormone receptors are nuclear proteins regulating gene expression, incorporating specific mRNAs known to be selectively affected by thyroid hormone would have been useful for this study. There are several genes known to be directly responsive to thyroid hormone during the perinatal period evaluated in this study, including:

MBP
PCP-2
RC3
Neurotrophins
Thyroid hormone receptors

In addition, there are several histological markers of thyroid hormone action in the developing cerebral and cerebellar cortex which would be potentially important, including size of the caudal portion of the corpus callosum, packing densities of cells in the cerebral and cerebellar cortex, and apoptosis in cerebellar granule cells. Finally, because compensatory responses of the fetus to hypothyroxinemia may be a very sensitive index of thyroid disruption, it would be important to incorporate measures of deiodinase activity and expression in these studies. Perchlorate treatment to the dams should be initiated at least two weeks prior to mating. In addition, perchlorate treatment should be continued throughout the period of lactation. Serum should be taken from a minimum of 10 dams at the time of birth and at the end of lactation.

Results

Circulating levels of thyroid hormones were evaluated to ensure that the perchlorate treatment produced effects observed in many previous studies. These measurements should be considered the foundation upon which all interpretion is made. Therefore, it is a major weakness of this study that thyroid hormone levels were found to be elevated by perchlorate treatment. This may be an artifact of low group sizes. Considering the importance of these data to the value of this study, it is unclear why so few animals were evaluated for thyroid hormone measurements. Measurements were taken in 3, 3, 3, 1, and 5 animals in each of the control, 0.1, 1.0, 3.0 and 10 mg/kg groups, respectively. However, this reviewer could identify only 4 dams marked as having serum taken for thyroid hormone measurements in the appendix provided -- it is not clear where the fifth value came from. These results jeopardize the relevance of the entire study. Since the doses of perchlorate used are well-known to reduce thyroid hormone levels, it is completely unclear how these hormone values could have been obtained. In addition, because perchlorate may be concentrated in milk, it would be important to measure thyroid hormone levels in pups. While the design called for hormone measurements in pups, no hormone levels are reported for pups.

It was found that the thickness of the corpus callosum and cerebellar cortex were slightly but significantly increased. These brain areas are known to be affected by thyroid hormone, but there is no evidence that these particular linear measures are related to perchlorate-induced effects on thyroid hormone action. These observations would not be related to reduced thyroid hormone action, which would have reduced the size of the caudal portion of the corpus callosum and a thinner cerebellar cortex. However, perchlorate is known to reduce circulating levels of thyroid hormone; thus, perchlorate would be predicted to reduce the size of the corpus callosum and the cerebellar cortex. If these measures are relevant to thyroid hormone, they confirm that the study did not produce known effects on the thyroid.

The summary of the neuropathology report suggests that observed changes were not considered biologically important without coincident histological abnormalities. This is a fundamentally illogical conclusion. Several neurological lesions caused by thyroid dysfunction are not correlated with specific histological abnormalities. For example, in the so-called TRß knock-out mice, homozygous animals cannot hear but the histology of their inner ear, at the light microscopic level, is not affected . In addition, hypothyroidism increases the rate of apoptosis in cerebellar granule cells 4-fold , but this only occurs in the internal granule layer, and only from postnatal day 2 to posnatal day 12 with a peak of apoptosis occuring on postnatal day 8. It is highly unlikely that a course end-point of "cerebellar thickness" measured from the roof of the fourth ventricle, would identify these important physiological changes especially if perchlorate-induced hypothyroidism is neither severe nor persistant.

There were several datasets in which a dose-response was not observed (e.g., TSH and thyroid hormone levels). However, because the interaction of perchlorate with the symporter will follow saturation kinetics, it is not unreasonable to expect that the effects on perchlorate will be saturable. This will be especially true in pups derived from perchlorate-treated lactating dams in which perchlorate exposure to the pups could be several-fold higher than that of the dams for reasons discussed above.

SUMMARY

Despite a great deal of difficult and meticulous work, several elements of this study limit its relevance for use in risk assessment. First, the duration of perchlorate exposure was not designed to identify thyroid effects. Initiating perchlorate dosing at mating would not allow hypothyroid effects to develop during pregnancy. Terminating perchlorate dosing in the middle of laction would not allow the full extent of lactational exposure to develop. Second, not all of the end-point measures employed in this study are sensitive to thyroid hormone or thyroid disruption. Those end-points that are sensitive to thyroid hormone are not the most sensitive that could be chosen. Moreover, the methods used and statistical evaluations employed were not those used to determine thyroid

hormone sensitivity of those end-points. Thyroid hormone produces very specific effects on brain development, not global effects on all brain areas. These specific neuro-developmental effects are reflected in specific behavioral deficits. The behavioral end-points chosen were not uniformly sensitive to thyroid hormone and largely cannot be used to identify thyroid-disrupting effects. Because perchlorate is believed to act on development by affecting thyroid hormone synthesis, it is unclear why the end-points were not tailored to identify thyroid end-points. It is recommended that gene expression be incorporated into studies on thyroid-disrupting compounds . Finally, failure to observe hypothyroxinemia in perchlorate-treated animals draws into question the efficacy of the treatments, and may invalidate this study.

Review of Neurobehavioral Study-Rabbit

Prepared by R. Thomas Zoeller, Ph.D.
University of Massachusetts
Department of Biology
Amherst, MA 01003

Review of Neurobehavioral Study-Rabbit

1. Design The purpose of this study was to provide information for the selection of dosages to be used in the developmental toxicity study of ammonium perchlorate in rabbits. This study incorporated the use of 25 rabbits: 5 groups of 5 rabbits each. Each group was given one of five target doses of ammonium perchlorate in drinking water (0, 0.1, 1.0, 10.0, or 20.0 mg/kg). This protocol was initiated on GD 6 through 28. In the two low-dose perchlorate groups, the doses were changed to 50 or 100 mg/kg on DG 13 and continued until GD28. The animals were terminated on GD29; blood was collected and the animals were examined for a number of gross physiological symptoms. The thyroid gland was isolated, prepared for histology and evaluated qualitatively.

The average T_3, T_4 and TSH levels were reported, but this reviewer did not see the raw data or the variance, so it is not possible to evaluate these results. No significant differences in circulating thyroid hormone levels were reported; however, base-line levels of T_4 and T_3 are quite different from those previously reported [Doohan, 1995 #1425] , though this may be a function of the strain of rabbit or other technical differences. However, no information is provided as to the validity of the assay for rabbits. The lab reports using a clinical assay kit which is a human serum-based assay. At the least, serial dilutions of rabbit serum should be evaluated for T_4 and T_3 to insure that there is linearity with the standard curve. This would determine whether there are factors in rabbit serum that interfere in some way with the assay. This is even more important for rabbit TSH which is likely to be immunologically different from human or rat.

There were treatment-related microscopic changes in the thyroid gland of rabbits o the 20, 50 and 100 mg/kg/day dosage groups. Treatment-related effects consisted primarily of folicular cell hypertrophy. These linear measurements may not be most sensitive to changes in thyroid hormone.

Review of Immunotoxicology Data

Prepared by R. Thomas Zoeller, Ph.D.
University of Massachusetts
Department of Biology
Amherst, MA 01003

Review of Immunotoxicology Data

Design

There was no statement of purpose in the description of the work accomplished ("Effects of Ammonium Perchlorate on Immunotoxicological, Hematological, and Thyroid Parameters in B6C3F1 Female Mice") or in the Review Document.

There were 20 different studies included in this report. In each study, female B6C3F1 mice, age 8-9 weeks, were exposed to 5 doses of ammonium perchlorate in the drinking water (0, 0.1, 1.0, 3.0 and 30.0 mg/kg/day). There were 6 animals per group for a total of 30 animals in each experiment. These 20 experiments fell into 8 categories:

1. Exposure to AP for 14 days and sacrificed
2. Exposure to AP for 90 days and sacrificed
3. Exposure to AP for 90 days, sacrificed at 120 days.
4. Exposure to AP for 14 days, *Listeria* challenge on day 7.
5. Exposure to AP for 90 days, *Listeria* challenge on day 86.
6. Exposure to AP for 90 days, B16F10 tumor cells injected on day 76.
7. Exposure to AP for 90 days, P815 tumor cells injected on day 79.
8. Exposure to AP for 14 days, P815 tumor cells injected on day 3.

The commission of these studies appeared to be motivated by several reports of immune compromise in patients receiving sodium perchlorate for the treatment of Grave's disease [Wolff, 1998 #1416]. This reviewer could find no information linking perchlorate or thyroid hormone to the various end-points used in these studies other than circulating levels of thyroid hormone. Because perchlorate is known to affect iodide uptake into the thyroid and potentially reduce thyroid hormone synthesis, it may be relevant to determine the effects of thyroid hormone itself on these immunological end-points.

Results

In two of the 90-day experiments, circulating levels of T_4 were significantly reduced by perchlorate. In one, T_4 levels were reduced in animals exposed to 1.0, 3.0 or 30.0 mg/kg/day. In the other, T_4 levels were significantly reduced only in

the 30.0 mg/kg/day group. The coefficient of variation in the control groups was nearly 30% in the latter experiment ("A") which probably accounts for the lack of statistical differences amoung groups. The T_4 assay was a total T_4 kit based on canine serum. No data are presented to validate this assay for mouse. Minimally, dilutions of mouse serum should be run to insure that interfering factors are not present in mouse serum.

Serum TSH was not affected by perchlorate in 14-day, 90-day, or 120-day studies. Again, the CV was large (30%-50%). The assay was a kit purchased from Amersham. This appears to be a rat TSH assay and while this assay may be valid for mouse, there is no information on validation of this herologous assay. This may be more important for TSH because the immunogenicity of this dimeric glycoprotein hormone is quite different among species.

The most consistent finding among the immunological end-points was a reduction in phagocytosis in perchlorate-treated animals. Although this finding was obtained in both 14-day and 90-day experiments, it was not repeated.

Critique

There is little if any background information on the effect of thyroid hormone on the immunological end-points evaluated in these studies. Likewise, there is no information on the effect of perchlorate on these end-points. Without this background information, it is difficult to fully interpret the present findings. The observation that perchlorate produced variable effects on circulating levels of thyroxine and TSH also is difficult to fully interpret without information about the characteristics of this heterologous assay.

Review of Toxicological Review Document

Prepared by R. Thomas Zoeller, Ph.D.
University of Massachusetts
Department of Biology
Amherst, MA 01003

1. In general, the review document appears to do a good job of capturing the key aspects of the protocols and conduct of the studies reviewed. The discussion appears to be balanced, providing a good sense of the strengths and weaknesses of the various studies.
2. In general, the review document appears to adequately evaluate the entire database. Specific issues of weaknesses both in the review document and in the design of studies are discussed below.
3. As written, the Review document provides new information on analysis of data presented from the various studies. These new analyses appear reasonable and informative and help improve the relevance of the studies performed.

General Issues

1. Several aspects of the the regulation of thyroid hormone "economy" are not fully presented throughout this document, and this weakness is reflected in the construction of experiments in the various studies. These are listed below:
    a. Page E-4, line 13: "differences in plasma protein binding...." It is true that the half-life of thyroid hormones in circulation are different among rats and humans. However, it is probably not true that this is simply due to differences in thyroid hormone binding proteins. First, rats do possess the same thyroid hormone binding proteins as expressed in humans (e.g., thyroid binding globulin, TBG, and transthyretin, TTR, [Tani, 1994 #1431]). However, these circulating proteins are regulated differently; however, pregnant and lactating rats, and neonatal pups, all express TBG in the circulation [Vranckx, 1994 #1432].
    b. Nearly 80% of circulating T_3 is derived from tissue metabolism of T_4. Moreover, tissue metabolism of thyroid hormones is dependent upon iodothyronine deiodinase activities which are themselves regulated by circulating levels of thyroid hormones [Burmeister, 1997 #1332; Germain, 1997 #1351]. Therefore, it is quite important to evaluate the potential effects of perchlorate on tissue metabolism of thyroid hromones, because these effects could be masking effects of perchlorate on the thyroid system.
2. The effect of perchlorate on the thyroid gland is the basis for focusing perchlorate studies on thyroid function and thyroid hormone action. This concept is well-developed in the review document and in the individual studies. However, there is no background information on the role of thyroid hormone in the development of end-points used for any of the studies. This is a critical issue because thyroid hormone has very specific, and in some cases subtle, effects on the development of the nervous system and end-points must be designed with this in mind to be valid. At the least, there is no discussion of

end-points that represent valid and sensitive measures of thyroid disruption during development. Moreover, most of the end-points used do not appear to be sensitive to thyroid disruption and thus, would not be expected to be affected by perchlorate.

3. In disussions of the physiological effects of perchlorate, the document acknowleges that effects are likely to be selectively mediated by the sodium/iodide symporter. Specifically, perchlorate blocks iodide uptake by the symporter. Moreover, the document acknowledges that the symporter is expressed in several regions - not just the thyroid gland - including choriod plexus, gastric and intestinal mucosae, and mammary tissue. What does not appear to be recognized in the review report or in the design of the neurodevelopment studies is the possibility (liklihood), that perchlorate blocks iodide uptake into milk and into the gut and thyroid of the pup. Moreover, it does not appear to recognize the possibility that perchlorate may itself be selectively taken up into milk and transferred to the pup. This is most important in the design of the neurodevelopment studies. However, this issue should be addressed in the review report especially with respect to the timing and duration of perchlorate exposure.

4. In discussions of perchlorate on thyroid hormone synthesis, there appears to be little recognition for the fact that the effect of perchlorate on thyroid hormone synthesis will take much longer to develop than the effect on iodide uptake. Because iodinated thyroglobulin is stored in the colloid, a considerable amount of thyroid hormone precursor is stored. Thus, even if doses of perchlorate are used which completely inhibit iodide uptake, it will still take some time for thyroid hormone levels to respond. This is especially important in the design of the neurodevelopment studies and their interpretation where perchlorate is initiated on the day of conception. In fact, hypothyroxinemia would not be predicted to occur until much later, perhaps as late as parturition.

Specific comments

1. P3-3, line 10, "This study suggested a LOAEL of 9 mg/kg-day in humans for short-term exposures." This dose is equivalent to 630 mg/day. This seems like a very large dose for a LOAEL when 9.7 mg/kg completely blocks iodide uptake.

2. P3-6, line 1: "There is no information to suggest that humans without Graves' disease would have a similar reaction to perchlorate". It is also true that there is no information to suggest that Graves' patients do not have a similar reaction.

3. Page 4-9, Line 26: "...there is a high-affinity binding protein, thyroxine-binding blobulin, which binds T_4 (and T_3 to a lesser degree); this protein is missing in rodents and lower vertebrates." This statement is false. [Tani, 1994 #1431; Vranckx, 1994 #1432].

4. Is the Review document as currently written useful for the puspose of characterizing the human health effects of ammonium perchlorate and the perchlorate ion?

   The database on perchlorate as described in this review document does not provide a great deal of confidence in establishing specific human health effects in a developmental context.

5. Development of Reference Dose (RfD)

   1. The reported NOAELs/LOAELs accurately reflect the findings as reported within the various studies. However, because many of the end-points were not the

most senstive to thyroid disruption, or analysis methods not the most sensitive, these values are not predicted to be the most accurate in fact. As an important example, the quantitative analysis of thyroid histology performed on pups in the Neurodevelopment study were not taken in rigorous manner. Therefore, this LOAEL only reflects the fact that the lowest dose of perchlorate produced an observed adverse effect. If perchlorate treatment were initiated prior to mating, and if the quantitation of thyroid histopathology were performed more rigorously, the LOAEL would undoubtedly be lower than 0.1 mg/kg-day.

2. The EPA's policy appears to be reasonable and it is appropriate to use the most sensitive end-point to base the RfD. In this case, using the thyroid histopathology study for pups is reasonable because this is a critical and sensitive time for thyroid hormone action and disruption.

3. The approach is reasonable as is the identification of the neurodevelopment study. It is important to emphasize that thyroid hormone actions on brain development represent irreversible effects; thus, the developing brain is essential to protect because there is no therapeutic remedy. However, for reasons discussed above and in the Neurodevelopment review, the estimate of 0.1 mg/kg-day is probably high.

4. Regarding the uncertainty factor of 3 for deriving the RfD: the greatest uncertainty that remains to be determined is the uptake of perchlorate in milk and the potential for concentration of perchlorate in the pup that occurs during lactation. If the assumptions about the transfer of perchlorate to the pup is correct, and there is good evidence that it is not, then the uncertainty factor of 3 is reasonable. However, in the absence of these data, an UF of 10 seems jusified.

5. The UF of 3 for using the LOAEL should be increased to 10 because of the issues discussed above for the estimation of this particular LOAEL.

6. The RfD calculated by the EPA and its justification is quite reasonable. However, the weaknesses in study design and data acquisition as described above introduces caution in this RfD.

IV. Further testing Needs for Perchlorate

1. It is essential to determine the concentration of perchlorate in milk of animals treated with different doses of perchlorate.

2. End-points need to be specifically designed to evaluate thyroid disruption. These are described in detail in the Neurodevelopment Review.

3. Although it is clear that perchlorate is a thyroid disruptor, the majority of end-points were not designed to identify thyroid disruption. This is critical especially considering the potential irreversible and specific effects of thyroid disruption during development. Thus, for an RfD to be developed with confidence, studies on perchlorate-induced disruption of thyroid hormone action is essential and these additional refinements would be important.

Review of 90-Day Subchronic Oral Bioassay Study

Prepared by Susan P. Porterfield, Ph.D.
CB-1104
Medical College of Georgia
Augusta, GA 30912

Submitted for the Peer Review workshop on the EPA Draft Toxicological Review Document on Perchlorate.

90-Day Subchronic Oral Bioassay Study Manuscript by Siglin et al., 1998

The experiments were designed to determine, in young adult rats, the oral RfD for ammonium perchlorate exposure by evaluating the effects of repeated exposure over 14-90 days with a 30 day recovery period. They looked for effects on the thyroid as well as numerous other organs. In addition, they evaluated the effects on select indices of reproductive function as well as assessing bone marrow micronuclei formation. Both sexes were analyzed separately.

1. Strengths and weaknesses of the experimental design

Strengths: The experiments represent an exceptionally thorough study of the toxicity of ammonium perchlorate (AP) in Sprague Dawley rats. They are studying the effects of AP in young adult animals over a 120 day period beginning at approximately 8 weeks postpartum. This period encompasses the late pubertal, pubertal, post-pubertal and young adult periods. They assessed thyroid function by evaluating serum total T4, T3, and TSH, thyroid size and histology. While this is far from a thorough evaluation of thyroid function, the end-points selected are appropriate for the screening desired based on preceding work by Caldwell et al., 1995. The dosing range was selected to encompass the appropriate ranges suggested in the work of Caldwell (0.01-10.0 mg/kg/day). In addition to the thyroid assays, they assessed reproductive status (estrus, testicular weights, sperm morphology, concentration, count, and motility). They evaluated body and tissue weight changes, food and water consumption, ophthalmology, hematology, routine clinical chemistry and bone marrow micronuclei formation. Males and females were evaluated separately because data of Caldwell et al. showed a gender effect. This gender effect was confirmed. A major strength of the project is the diversity of the parameters studied.

Weaknesses: The sample sizes were 10 animals/sex/treatment/age. For some analyses, the sample size was so small relative to the variability that it made it impossible to conclude that no effect existed. This is particularly notable in some of the serology studies.

Data for some of the serology were not within the range expected from historical data, even with the controls. Examples include the following (numbers in parentheses are the historical data):

- Bun 2.9 ng/dL (13.7)
- Creatinine 0.039 ng/dL (0.428)
- Glucose 299 mg/dL (104)
- Potassium 6.9 mMol/L (4.8)
- Calcium 13.47 mg/dL (10.38)
- Cholesterol 53 mg/dL (35.2)
- Phosphorus 13.15 (8.46)
- AST 37.7 IU/L (92.1)
- ALT 63 IU/L (34.7)

There was considerable variation in weight of the animals at the beginning of the study (139-173 females; 169-222 males). Because much of the data were expressed as, or on the basis of weights, the variability was increased by the diversity of initial weights. The animals were randomly assigned to treatment groups.

In some instances, the small sample size (10) increased the likelihood of a type II error because the SD was large relative to the mean. For example, (Table 7) monocytes (control mean = 0.2; SD = 0.17; eosinophils mean = 0.1; SD =0.09). Could the data be reported with more significant figures? Because of the suggestions that perchlorates could cause leukopenia or agranulocytosis in humans, these data are important. In table 8, the value reported for sorbitol dehydrogenase in males at day 92-95 with the 0.01 dose is 3x the average dose for other treatments and ages. Is there any explanation for this variance?

The data for semen analyses are difficult to interpret because the high SDs increase the likelihood of a type II error with the small sample size used. More animals would have improved this portion of the study.

In some cases, there were large differences between control values of TSH, T4, and T3 between ages such as 90 and 120 days when large differences would not be anticipated. For example female T3 is 4.4 at 90 days and 3.5 at 120 days. Were this value comparable to the 4.4 seen at 90 days, a treatment effect would be seen during the recovery period. Serum T3 (female) was 170 at 90 days and 224 at 120 days. Were the 120 day value as high as the 90 day value, no recovery would have been seen. The variability among the controls is not uncommon with hormone assays and represents a reason for a larger sample size.

Are the questions clearly identified? Yes

Are they important? Yes

Is the study design appropriate to answer the questions? The only problem I see with the study design is the small sample size relative to the anticipated variability in some of the assays and the variability of initial body weights for the animals used in the study. While the thyroid assays do not represent a definitive assessment of thyroid function, they do include the most appropriate parameters for screening.
Limitations? The major limitation is the increased risk of a type II error in the assay systems where the variability is high.

2. Limitations in the conduct of the study
The major concern is the large deviation from historic values seen in many of the serologic chemistry evaluations. Fed control rats would not be expected to have serum glucose levels approaching 300 mg/dL.

3. Statistics
The statistical analyses appear appropriate for the study design.

4. Strengths and weaknesses of the presentations in the study report
The investigators provided extensive tables documenting in detail the experimental results. While the positive data support the conclusions drawn, I have reservations about the likelihood of type II errors with some of the data. Consequently, I have difficulty with their conclusions that no treatment effects are seen with respect to some of the parameters studied.

5 & 6. Was the study of sufficient quality for use for hazard characterization purposes?
I have concerns that the serologic and hormone data sets need more animals for validation. While the levels showing effects are meaningful, I question the validity of conclusions about those treatment levels where no significant effects are seen. An additional component of the experiment that could be repeated before concluding that there are no effects is the sperm analyses.

REVIEW OF TOXICOLOGICAL REVIEW DOCUMENT

A. Effects of concern to human health

The review provided a thorough assessment of the available data. While time precludes my having studied all experiments leading up to the review document, those that I have reviewed are appropriately represented. The additional statistical analyses are often appropriate and clarify significant points. There are several minor corrections I would note in the report.

p. 4-16, lines 10-12. While thyroid hormone nuclear receptors are thought to be the predominant mechanism of action, most investigators do not consider them to represent the only mode of action of thyroid hormones.

p. 4-16, lines 15-17. The paper by Cao is appropriately cited when the authors suggest that the reason for the increased severity of iodine deficiency relative to congenital hypothyroidism is due to iodine deficiency per se rather than the effect of the iodine deficiency on thyroid function. However, most investigators working in this field would not agree with that conclusion and would suggest that the increased severity and irreversibility is a result of hypothyroidism beginning at the time of conception versus beginning in mid-gestation. In addition, maternal thyroid function is an important determinant of fetal/placental nutrition.

p. 4-17, line 19. Luteinizing hormone is the correct spelling rather than leutenizing.

p. 5-11, figure 5-2, figure 5-4. The legend includes male and female but the data were collapsed. The legend needs to be removed.

Tables 6-1-A-6-2 were particularly helpful.

III. Hazard Characterization

I am not qualified to assess this

Review of Developmental Studies

Prepared by Rochelle W. Tyl, Ph.D.
Research Triangle Institute
Center for Life Sciences and Toxicology
Research Triangle Park, NC 27709

I. REVIEW OF INDIVIDUAL STUDIES INITIATED SINCE MAY 1997

A. Oral (Drinking Water) Dosage-Range Developmental Toxicity Study of Ammonium Perchlorate in Rabbits. Final Pilot Report; Protocol No.: 1416-002P. R.G. York, Study Director, Argus Research Laboratories, Inc., Horsham, PA. Report Date: December 10, 1998.

1. Study Design

This is a relatively standard study design for a range-finding study to select target doses for a subsequent "definitive" developmental toxicity study. The design involved five naturally-mated does (presumed pregnant) per group (five groups) to achieve target intakes of ammonium perchlorate of 0, 0.1, 1.0, 10.0, and 20.0 mg/kg/day, in groups 1 through 5, respectively (with concentrations in the drinking water of "0, 0.96, 9.75, 95.75, and 191 µg/ml", page 6), based on body weights on gestational day (gd) 5 and estimated water consumption of 100 ml/kg/day, with subsequent adjustments based on actual body weights and water consumption values. Correcting the water concentrations, based on the previous week's body weight and actual water consumption values, will always be off (too low) since the performing laboratory is basing the subsequent week's concentrations on the previous week's body weights, and the animals are gaining weight during the pregnancy. On gd 13, the target concentrations were increased in groups 2 (0.1 mg/kg/day) and 3 (1.0 mg/kg/day) to provide target intakes of 50 and 100 mg/kg/day, respectively, "in order to establish evidence of maternal toxicity." The protocol (Attachment 1) and the text of the report do not provide information on the statistical analyses used. The *ad libitum* exposure period to the does was for gd 6-28, with scheduled sacrifice on gd 29. At sacrifice, the doses were examined for body weight, gravid uterine weight, fixed thyroid (and parathyroid) weight, and gross lesions of organs and body cavities. Maternal blood was taken at sacrifice (from the vena cava) for analyses of TSH, T3, and T4 (by AniLytics, Inc., Gaithersburg, MD), and the fixed thyroids were examined histologically by Research Pathology Services, Inc. (New Britain, PA). The number of ovarian corpora lutea and the number and status of uterine implantation sites (total, early, and late resorptions, live and dead fetuses) were recorded. Fetuses were counted, weighed, examined for gross (external) alterations, and sexed internally (one cannot sex rabbit fetuses externally). Fetal internal (visceral) examinations were not performed (except for sex determination), and fetal thyroids were not examined.

2. Conduct of Study

The study was performed under EPA (FIFRA/TSCA) GLPs with a QA statement provided (Attachment 4). The changes in target concentrations and intakes in groups 2 and 3 on gd 13 were done to provide higher intakes to ascertain maternal toxicity (obviously, to that point, the performing laboratory did not see any evidence of maternal toxicity through 20.0 mg/kg/day). This change was documented by Amendment No. 1, Attachment 1, dated December 30, 1997, but the amendment did not specify the calendar date(s) and gd that the change was initiated. (Hopefully, this information is in the study records.) This change makes interpretation of fetal findings at term very problematic, since embryos in groups 2 and 3 were potentially exposed to low doses early in major organogenesis (gd 6-13) and to very high doses (in a "hyperthyroid" mother) for the rest of the embryonic (organogenesis) period and throughout the fetal period (see discussion in item no. 5 below). Fetal thyroids should have been examined. The performing laboratory removed one doe at 0 mg/kg/day from study since she had a one-horn pregnancy with only three live fetuses. This is reasonably appropriate (and should be specified in the laboratory's SOPs).

3. Statistical Methodology

Neither the protocol nor the report specifies which statistical tests were used. The protocol (p. 12) only states "Averages and percentages will be calculated. Litter values will be used where appropriate. Additional procedures and/or analyses may be performed if deemed appropriate." This provides no useful information. I am assuming that nonparametric statistics were employed for continuous (body weights and changes, feed and water consumption, etc.) and discrete (malformation incidences, etc.) data because of the small group size (4-5 does per group), with a $p \leq 0.05$ as standard for statistical significance from the concurrent control group values. The summary and individual animal data tables also do not specify which statistical tests were used for which parameters.

4. Presentation of Results

These were adequate, except there was no indication of which statistical tests were performed on which parameters. Three fetuses in three different litters (out of five, 60% litter incidence) at 20.0 mg/kg/day exhibited major external malformations. These fetuses were in the group which was exposed to the highest dose during the first half of major organogenesis (the 50 and 100 mg/kg/day groups were not switched to the higher doses until gd 13). Based on the performing laboratory's historical control data (Appendix J to the definitive study final report), meningocele and umbilical hernia (in this study in male fetus 6747-3) were present in low incidences (litters 0.85%, range 0-50%, fetuses 0.11%, range 0-12.0%, for umbilical hernia; and litter 0.24%, range 0-16.7%, fetuses 0.03%, range 0-2.0% for meningocele). Female fetus 6748-3 had cleft palate and a number of other major malformations (only cleft palate in historical control, litter 0.12%, range 0-5.3%; fetuses 0.01%, range 0-0.6%). Male fetus 6749-1

exhibited distended abdomen (ascites), not in historical controls. The performing laboratory acknowledges this could be due to early organogenesis exposure or chance, and correctly suggests that this dose be employed in the definitive study. My copy is missing page 2 and following of Attachment 2, "Thyroid Hormone Levels" from Anilytics, Inc., so I cannot examine individual data for groups 3 through 5.

5. Use in Hazard Characterization

This study was designed, performed, and interpreted to assign target doses (concentrations in the drinking water and intake values) for the subsequent definitive developmental toxicity study. It served this purpose. The changes in doses on gd 13 means that the does (and the embryos via possible transplacental transport) were exposed to low intakes (0.1 and 1.0 mg/kg/day) for the first half of the period of major organogenesis (gd 6-13), and the does (and embryos/fetuses) were exposed to high intakes (50 and 100 mg/kg/day) for the second half of major organogenesis (gd 13-18/19) and throughout the fetal period (gd 18/19-28). Since the fetal thyroid, at least in humans, begins to concentrate iodide in the second and third trimesters (the fetal period in humans), the embryos/fetuses in these groups potentially were exposed during a critical period of differentiation and prenatal function to high levels of perchlorate. Examination of their thyroids (at least) would have been informative.

6. Incomplete Study

The study is complete, and the final report is available.

B. Oral (Drinking Water) Developmental Toxicity Study of Ammonium Perchlorate in Rabbits. Final Report; Protocol No. 1416-002. R.G. York, Study Director, Argus Research Laboratories, Inc., Horsham, PA. Report Date: September 1, 1998 (final report), September 10, 1998 (report amendment).

1. Study Design

This is a relatively standard study design for a Segment II developmental toxicity study in rabbits. The study design involved 25 naturally-mated does (presumed pregnant) per group (six groups) exposed to the test material *ad libitum* in the drinking water, from gd 6 through 28, to achieve target intakes of ammonium perchlorate of 0, 0.1, 1.0, 10.0, 30.0, and 100.0 mg/kg/day, with adjustments to the concentration of test material based on the previous week's body weights and water consumption. Correcting the water concentrations based on the previous week's body weights and actual water consumption values will always be off (too low), since the performing laboratory is basing the subsequent week's concentrations on the previous week's body weights, and the animals are gaining weight during the pregnancy. Two of the target intakes in this study (10.0 and 30.0 mg/kg/day) bracketed the dose in the range-finding study (20 mg/kg/day) at which fetal malformations were observed. Does were examined and weighed, and feed and water consumption was measured daily. Maternal rabbits were sacrificed

on gd 29, with blood taken from the inferior vena cava for subsequent analysis for T3, T4, and TSH by AniLytics, Inc. (Gaithersburg, MD). Body weights, gravid uterine weights, and fixed thyroids (plus parathyroids) were weighed, and the thyroids were evaluated histologically by Research Pathology Services, Inc. (New Britain, PA). The number of ovarian corpora lutea and the number and status of uterine implantation sites (total, early and late resorptions, dead and live fetuses) were recorded. Each fetus was weighed, examined grossly, and sexed internally. Live fetuses were euthanized. For approximately 50% of the fetuses per litter, a single cross section was made between the parental and frontal bones of the skull, and the brain was examined *in situ*. The remaining fetuses per litter were decapitated and the heads fixed and decalcified in Bouin's solution for subsequent free-hand serial sections and examination. All fetuses were examined viscerally (abdominal and thoracic organs) and skeletally after staining with alizarin red S (for ossified bone). The report indicates that the skeletal preparations were examined for "skeletal and cartilaginous alterations." A deviation from the protocol and SOPs (Appendix D) indicates that although the protocol and SOPs specify staining with alizarin red S and Alcian blue (the latter specifically for cartilage), "The skeletal specimens were stained with only alizarin red S, not also with Alcian blue." The deviation continues "This deviation did not adversely affect the outcome or interpretation of the study, because although staining with Alcian blue makes it easier to see the cartilage, it is not necessary to stain the cartilage to evaluate it. Bones and cartilage are evaluated adequately using the single stain, and double staining is not a guideline requirement." This reviewer has two major problems with this deviation: (1) it is extremely difficult, if not impossible, to "adequately" assess ossified and cartilaginous skeletal components without double staining (tricks to enhance visualization of cartilage in single-stained specimens include bottom lighting and use of ethanol to make the cartilage "cloudy", but these are <u>not</u> adequate); and (2) if the performing laboratory violated the study protocol and their own SOPs (presumably an inadvertent omission; did they "forget"?) for this important procedure, what else did they forget? Fetal thyroids were not weighed or assessed histopathologically, and fetal blood was not evaluated for TSH, T3, or T4.

2. Conduct of Study

This study was performed under EPA (TSCA/FIFRA) GLPs, and there is a QA statement (Appendix K). See comments in answer to question 1 on the major deviation for the protocol on staining fetal skeletal specimens. The lack of staining for fetal cartilaginous skeletal components impairs the laboratory's ability to adequately assess the fetal skeleton at term and identify any possible treatment-related changes.

3. Statistical Methodology

It is adequately described in the protocol (Appendix C, page 14) and in the final report (pages 28-29). However, I have two concerns:

a. The decision tree provided uses the Bartlett's test for homogeneity of variances to decide whether to use parametric or nonparametric statistics. If Bartlett's is significant for a given parameter at $p \leq 0.05$, nonparametric analysis is performed. This sets the "gate" for parametric versus nonparametric analyses very (too) low. Most laboratories, including this reviewer's, use a $p<0.001$ as the gate. I went back and looked at our rabbit data; if we had used the cutoff of $p<0.05$ for Bartlett's, almost all parameters would have been examined nonparametrically. Nonparametric statistics are less robust than parametric statistics, so fewer differences would be identified as statistically significantly different using nonparametric statistics. In my laboratory, we perform an arcsine-square root transformation on all litter-derived percentage data to allow use of parametric methods. For these litter-derived parametric data, we also weigh the ANOVA according to litter size (litters with a larger number of fetuses have lower weight fetuses, relative to litters with fewer numbers of fetuses).

b. Neither the text, summary tables, nor individual tables indicate which tests were used to evaluate statistical significance, so the reviewer has no way of knowing whether parametric or nonparametric tests were used.

4. Presentation of Results

Adequate, although the specific statistical test used is not provided in text or summary tables for a given parameter (see answer to #3).

5. Use in Hazard Characterization

This study is useful and appropriate to use in hazard characterization. The intake values must be corrected for presence of the ammonium ion, and fetal thyroids and hormone status (a potentially sensitive life stage) were not evaluated.

Question: Was a rat developmental toxicity evaluation performed (usually two species, rat and rabbit, are used for developmental toxicity assessments)? Since the rat is apparently as sensitive as or more sensitive than the human, this would be a critical study (or an important data gap if there is not one available).

6. Incomplete Study

The study is complete, and the final report is available.

Review of Two-Generation Reproductive Study

C. Oral (Drinking Water) Two-Generation (One Litter per Generation) Reproduction Study of Ammonium Perchlorate in Rats. Final Interim Report, Protocol No. 416-001. R.G. York, Study Director. Argus Research Laboratories, Inc., Horsham, PA. Report Date: September 15, 1998.

1. Study Design

This is a two-generation study in CD® (Sprague-Dawley) rats, one litter per generation, performed according to U.S. EPA FIFRA "Pesticide Assessment Guidelines Subdivision F, 83-3" (this is incorrect; it should be 83-4), with additions from the OPPTS draft guidelines (1996), and thyroid (and parathyroid) weights and histopathology for parental and weanling animals. Blood was also taken from parental and weanling animals for "possible analysis of TSH, T3, and T4" (report, p. 17). The study design involved 30 rats/sex/group (four groups) exposed *ad libitum* (summary, p. 16, is incorrect; it states "once daily") to ammonium perchlorate in the drinking water, to achieve target intakes of 0, 0.3, 3.0, and 30.0 mg/kg/day. Concentrations in the drinking water were adjusted based on the previous week's actual body weight and water consumption values. Correcting the water concentrations, based on the previous week's body weight and actual water consumption values, will always be off (too low), since the performing laboratory is basing the subsequent week's concentrations on the previous week's body weights and the animals are gaining weight rapidly during the F0 and F1 prebreed exposure period and during the F0 and F1 gestational periods. There was also no difference in the concentration provided to the males versus the females, but they differ in their body weights and water consumption and will, therefore, differ (with a fixed concentration the same for both sexes) in their intake in mg/kg/day. There was a ten-week prebreed exposure period (with estrous cyclicity evaluated for the last three weeks), a two-week mating period (based on examination of report Figure 1; not specified in report), three-week gestation for F0 and F1 parents, and three-week lactation period for parents and F1 and F2 offspring. F1 offspring will be evaluated for acquisition of preputial separation (males) beginning on postnatal day (pnd) 39 and of vaginal patency (females) beginning on pnd 28 (too late; CD® rats begin to acquire vaginal patency prior to pnd 28). The current report is an INTERIM report through the necropsy of F1 weanlings; there is no analytical report (Appendix G), no histology report (Appendix H), and no indication that parental and offspring blood will be analyzed for TSH, T3, or T4.

2. Conduct of Study

Based on the protocol and data available, it is a well-conducted study.

3. Statistical Methodology

Statistical analysis is described in the report (pages 37-38) and in the protocol (Appendix D, page 23). I have five concerns with the methodology as presented:

a. It does not provide adequate detail.

b. The use of Bartlett's test for homogeneity of variances to decide whether to use parametric or nonparametric tests with the "gate" at $p<0.05$ (see comments to the rabbit developmental toxicity study) and the implications for robustness of analyses.

c. There is no mention of analysis of covariance (ANCOVA) for day (age) of acquisition of preputial separation (F1 males) or of vaginal patency (females) with body weight at acquisition as the covariate. A number of laboratories running multigeneration studies under the new OPPTS guidelines have data indicating that acquisition of these developmental landmarks is at least partially dependant on body weight, so that lighter pups (due for example to systemic toxicity) will acquire later.

d. There is no mention of analysis of covariance (ANCOVA) by body weight for anogenital distance at birth of F2 pups if triggered by alterations in reproductive development in F1 animals. Again, anogenital distance is sensitive to body weight.

e. There is no indication of which tests were used to assess statistical significance for a given parameter in the summary tables.

4. Presentation of Results

Based on the interim nature of this report, presentation is adequate. There should be an indication of which statistical test was used for each parameter, parametric or nonparametric.

5. Use in Hazard Characterization

When all the data through weaning of the F2 offspring are in, the thyroid weights and histopathology in parents and offspring, and hopefully parental and offspring TSH, T3, and T4 values, will prove sensitive endpoints to confirm (or refute) the current most sensitive LOAEL of 0.1 mg/kg/day for the pnd 5 pups in the developmental neurotoxicity study. The low dose in this study is 0.3 mg/kg/day, higher than the most sensitive LOAEL, but it will provide confirmation and confidence in the assessment of effects at 0.1 mg/kg/day in the other study.

6. Incomplete Study

This report is INTERIM and ends at the weaning of the F1 offspring. It does not include analytical results of the drinking water, histopathology of parental and offspring tissues, including thyroid/parathyroid, and it does not include the postnatal growth, development (including reproductive) and mating, and

pregnancy and lactational events for F1 parents and F2 offspring. The F1 generation is really the most important, since it was at least potentially exposed during gestation and lactation, with exposures continuing through its mating, gestation and lactation, to produce F2 offspring. Since the fetal hypothalamic-pituitary-thyroid axis may be at especial risk and there are known effects of hypothyroidism on growth, development, and reproduction, this generation may show effects at the low dose of 0.3 mg/kg/day. Since this dose is higher than the low dose for the developmental neurotoxicity study, 0.1 mg/kg/day, at which effects were observed in offspring on pnd 5, it will not alter the most sensitive LOAEL, but it will provide confidence in the assessment.

II. REVIEW OF TOXICOLOGICAL REVIEW DOCUMENT

A. Effects of Concern to Human Health

1. Key aspects of study design (protocol), conduct, and conclusions of each toxicology study have been adequately described. Limitations have been appropriately discussed.

2. The strengths of the analyses have been well described.
   i. The rat is a good surrogate for humans (rat is at least as sensitive as humans and probably more sensitive).
   ii. The concept of merging cancer and noncancer endpoints is appropriate, given the early and lower dose histopathologic changes to the thyroid and changes in thyroid hormone homeostasis (TSH, T3, T4) as health risks in and of themselves, and as a predictor ("harbinger") of possible carcinogenic risk is appropriate. Even in the absence of mutagenic potential of perchlorate, the pre- and perinatal animal (and human) should be considered at especial risk.
   iii. The discussion of the hypothalamic-pituitary-thyroid axis is appropriate.
   iv. The uncertainty factors are appropriately conservative and protective.

   The biggest weakness is the lack of a final report for the two-generation study (only an interim report is available). The F1 generation is crucial and potentially the most sensitive. However, the low dose is 0.3 mg/kg/day, so a lower NOAEL/LOAEL will not be identified.

3. The additional statistical analyses were useful and appropriate. The designation of 0.1 mg/kg/day as a minimal LOAEL, based on the pnd 5 thyroid effects for the developmental neurotoxicologic study, is appropriate. This is in keeping with biological plausibility, mode-of-action, weight of the evidence, and the "precautionary principle."

4. I don't know of any additional references.

B. NA (Ecotox)

C. Additional Issues

1. Under exposure characterization (e.g., page 7-8, line 1 and page 8-6, line 7), the document indicates that "0.37%" is equal to "37×10^6 µg/L." But, 0.37% = 3700 ppm (0.0037 [conversion from percent] x 1,000,000 [1 liter = 1,000,000 µl] = 3700 ppm). Therefore, the correct value must be 3.7×10^6 µg/L (3700 mg/L), at least based on my calculations, which is possible, given the solubility of perchlorate in water of 24.992 w/w % for the ammonium salt (page 2-3, Table 2-2).

2. The document is useful.

III. HAZARD CHARACTERIZATION

A. Development of Reference Dose (RfD)

1. The individual NOAELS and LOAELS for each study, based on EPA reanalysis, are appropriate.

2. The EPA approach, using a merged cancer and noncancer endpoint evaluation, and viewing changes in histopathology of the thyroid and in TSH, T3, and T4 as noncancer endpoints of concern and predictors ("harbingers") of possible cancer risk, are appropriate and justified.

3. The selection of 0.1 mg/kg/day as the "minimum" LOAEL, based on histopathological changes in the thyroid in pups on pnd 5 in the developmental neurotoxicology study, is appropriate.

4. Given the nonmutagenicity of perchlorate, the thyroid-pituitary disruptions in homeostasis (histopathology and circulating hormone levels), and the rat as an animal model at least as sensitive as humans (and apparently more sensitive), an uncertain factor of "3" is very conservative to go from rats to humans. A factor of "1" might be more realistic and still protective.

5. The additional uncertainty factor of "3" to go from a LOAEL to a NOAEL is appropriate, given that histopathologic changes are very sensitive and there were such changes, albeit "minimal" at 0.1 mg/kg/day in the pnd 5 pups.

6. The use of the RfD for a harmonized human health risk estimate should be protective for both noncancer health effects and cancer endpoints of perchlorate ion. The data support this position.

B. NA (Ecotox)

IV. FURTHER TESTING NEEDS FOR PERCHLORATE

A. Toxicological Testing

1. The experimental designs of the toxicology tests undertaken since May 1997 are adequate, except for:

   i. The rat multigeneration study low dose is 0.3 mg/kg/day, while the low dose and the subsequently identified "minimal" LOAEL in the developmental neurotoxicology study is 0.1 mg/kg/day. "Hindsight is 20-20," but it would have been very useful to run 0.1 mg/kg/day and a

lower dose in the multi-generation study to confirm or refute the 0.1 mg/kg/day LOAEL and identify a true NOAEL in a study examining offspring at least potentially exposed during gestation and lactation, with exposure continuing through their reproductive phase (the F1 to make F2 litters).

ii. I hope the blood samples retained from F0 and F1 parental animals and F1 and F2 weanlings for the multi-generation study WILL be analyzed for TSH, T3, and T4 to accompany the histopathological assessments on the thyroids from these animals.

2. If there has not been a Segment II developmental toxicity study of ammonium perchlorate in rats (the usual approach, mandated by EPA FIFRA and other regulatory testing requirements, is this test in two species, a rodent, usually rat, and a nonrodent, usually a rabbit), it SHOULD be done, since the rat is as sensitive as/or more sensitive than the human. *In utero* exposure, up until term, will assess effects of maternal "hypothyroidism" and possible embryo/fetal "hypothyroidism" on *in utero* development, especially in the postembryonic fetal period, which corresponds in humans to when the fetus begins to concentrate iodide into the thyroid (does this indicate activation of the symporter?). The protocol should be a standard OPPTS (1998) developmental toxicity study with maternal blood and maternal (and fetal?) thyroids taken for serum TSH, T3, T4 (dams), and histopathology (dams and fetuses).

3. Development of a PBPK model to address species differences in iodide uptake, perchlorate kinetics, and subsequent perforbations in the hypothalamic-pituitary-thyroid axis is an excellent idea, but:

   i. It will only be as good as the data available and will take considerable time to develop.
   ii. It must model maternal, embryo/fetal and perinatal offspring parameters, with more than one timepoint during the pregnancy to detect anticipated changes in the parameters of concern (e.g., iodide uptake, as they change in the maternal and embryo/fetal compartments during the pregnancy).

B. NA (Ecotox testing)

NOTE: There are a number of typographical errors and more substantive errors in the EPA Toxicological Review (review draft) which will need to be corrected prior to finalization. Examples (not complete) follow:

Review Page	Review Line	Error
E-6	12	Pimephales (fathead minnow) is an aquatic vertebrate (not invertebrate)
1-1	22-23	Thermal explosive decomposition occurs above (not below) 300°C.
3-2	23	"given no [not on] more than..."
3-6	17	"perchlorate" fix typo
3-6	19	"...adverse effect than healthy..."
3-7	17	"...radioimmunoassay..."
4-2	31	"from... to [not and] where..."
4-7	Table 4-1	Superscript "d" and "e" in table but not in footnotes; no superscripts "b" or "c" are present
4-8	8	"...quarternary...", fix typo
4-10	Figure 4-3	Item 9, "TSH-secretory..." [not THS]
4-11	Table 4-2	under "Indirect": (1) "- chemicals inhibiting TH [add "release"], and (2) "- chemical [add "s"] inhibiting [should be stimulating?] hepatic..."
4-18	6	"...developing conceptus..." not "fetus" (too narrow a term)
4-22	Table 4-6	middle column: (a) under rat dev. neurotox, maternal animals are "dams" not does; (b) designate initial parental generation "P0" (dev. neuro.) or "F0" (two-gen), but be consistent
5-43	13	Dev. Tox studies do not employ "brain histology," but serial freehand sections of fixed tissue (no embedment, microtome section or stain, etc.)
5-55	31	"affected" not "effected"
5-56	17	"perchlorate", fix typo

(continued)

Review Page	Review Line	Error
5-57	30	The description of what the Ames test measures is incorrect. It measures "the reversion from a histidine- (histidine dependent) state to a histidine+ (histidine independent) state induced by chemicals..." (by growing the exposed organism on media without histidine, only his+, independent revertants will grow)
5-59	29	"...drinking water gavage" is incorrect, can't be both; should be "gavage" (I think)
6-10	5	"...diurnal...", fix typo
6-12	9	"...trichloroacetate...", fix typo
6-30	29	"...distributions...", fix typo
6-32	14	"...homeostasis...", fix typo
6-32	21	"...or those treated with anti-thyroid drugs..."; move "anti-thyroid"
6-41	8	"...exquisite...", fix typo
6-45	19	"...because it fell..." (not "if")
6-48	17	"...appropriate...", fix typo
6-52	7	"... because of the efflux..." (delete "to")
7-8, 8-6	1, 7	see comment on "0.37% (37 x 106 μg/L)"
7-14	2-3	If survival is reduced as indicated, it is not dose related; should it read "survival was reduced by 26%...etc."?
9-6	48	"...N-bis (2-hydroxypropyl) nitrosamine...", fix typo

Review of Immunotoxicity Studies

Prepared by Kimber White, Ph.D.
Medical College of Virginia/Virginia Commonwealth
Strauss Immunotoxicology Research Laboratory
Richmond, VA 23298

I. REVIEW OF INDIVIDUAL STUDIES INITIATED SINCE MAY 1997

1. The evaluation of the effects of perchlorate on the immune system, as conducted in the immunotoxicological studies, contributes significantly to the characterization of perchlorate. The question being investigated in these studies, does perchlorate adversely affect the immune system, is certainly germane to the evaluation of the overall health effects of perchlorate exposure. Overall, the experimental design for the immunotoxicology studies appeared appropriate. One of the major strengths of the experimental design followed is the fact that the assays were conducted at multiple time points (14, 90, and 120 days) and, more importantly, each of the assays was conducted at least twice. This approach of repeating all functional assays at least twice is not routinely done in most immunotoxicological investigations, and the fact that the assays were repeated in the evaluation of perchlorate is considered to be a major strength of the study. A second significant strength is the inclusion of the 30-day recovery period (120 day results) in the experimental design. Most immunotoxicological evaluations do not include recovery information on the test compounds. The selection of the B6C3F1 mouse as the mouse strain for use in the immunological investigations is also considered a strength since significant data exist on the immunological responses of this test species. Furthermore, the use of sentinel mice during the study period also represents a strength since the immunological assays are very sensitive to viral infections.

The experimental design also had several weaknesses. First, the decision to use only 6 animals per group is considered to be a weakness. As reflected in the data, on several occasions multiple samples from the same group were lost as a result of technical error or for some other reason, resulting in small group sample size. A decrease in the number of samples in the control or treatment groups can have a significant impact on the statistical evaluation of the results. Another weakness in the design is the fact that a positive control was not included in each of the functional immunological assays. By having a positive control present in the assay, one insures that the assay was conducted correctly and that the assay was capable of detecting an effect, if one was present. The use of a positive control is also helpful as a reference point in the evaluating the effect the test compound has on an immunological parameter being measured.

Only one sex, female animals, was evaluated for immunological effects. While this does represent a limitation of the studies, it is consistent with the usual approach used by other organizations, such as the National Toxicology Program, in evaluating the effects of test

materials on the immune system. The decision to use female mice and not male mice is considered to be a strength in the design.

The major weakness in the experimental design was the selection of the immunological assays and host resistance assays used in the evaluation of perchlorate. It is unclear from the information provided what the rational was behind the selection of the assays. A tremendous amount of work has gone into developing, validating, and determining the sensitivity and predictability of various immunotoxicological assays (Luster *et al.*, 1988, 1992, and 1993). This work appears not to have been considered in the assay selection for the evaluation of perchlorate. For example, the most predictive immunotoxicological assay, the IgM antibody-forming cell response to the T-dependent antigen, sheep erythrocytes (Plaque Assay), was not originally undertaken in the evaluation of perchlorate. From the supplied material, it appears that this assay may have been added at the insistence of the EPA; however, no results were provided for evaluation. Some of the assays conducted in the evaluation of perchlorate are no longer used by other organizations since they have been found not to be very predictive of immunotoxic effects. In addition, the procedures followed in conducting some of the assays differ significantly from those followed by Luster *et al.* For example, routinely a Delayed Type Hypersensitivity (DTH) assay, such as those conducted in the Luster *et al.* studies, consists of a holistic *in vivo* evaluation of the response to the test article. In the Luster *et al.* DTH studies, a mouse or rat is both sensitized and challenged *in vivo* and the endpoint monitored occurs *in vivo*, such as the swelling of a foot pad or recruitment of cells to the site of challenge. The "DTH" evaluated in these studies is more accurately described as an antigen specific proliferative response as opposed to a "classical" *in vivo* DTH response. As such, it losses the predictive value for identifying potentially immunosuppressive compounds.

It is also unclear what was the rational for selecting and conducting the host resistance assays prior to the completion of the functional assays. If humoral immunity, as evaluated in the plaque assay, is affected to a much greater extent than cell-mediated immune parameters, such as the cytotoxic T-cell response, a host resistance assay designed to evaluate humoral immunity may prove to be more appropriate in the overall immunotoxicological evaluation of perchlorate than the *Listeria monocytogenes* assay which was conducted. Since effects were observed on natural killer cell activity, the selection of the B16F10 tumor was an appropriate model for evaluation of innate immunity host resistance.

2. The immunotoxicology studies conducted in the evaluation of perchlorate were reported to have been conducted under the guidelines of Good Laboratory Practices (GLPs). As such, the integrity of the data, from collection until inclusion of tables and graphs in the final report, would have been appropriately monitored by the Quality Assurance Manager and found free of error. Furthermore, having conducted the study under GLPs indicates that all Standard Operating Procedures (SOPs) for the various assays conducted were in place and were followed in carrying out the study.

A limitation which could decrease the relevance of the study findings is the technical problems with some of the assays. On several occasions the studies' results did not appear to repeat. This may just be due to the fact that the assays were conducted using living biological systems. However, technical problems can also contribute to the failure of an assay to produce reproducible results. The report authors themselves indicated technical problems occurred in obtaining some of the data (i.e., spleen and thymus cellularity, nitrite production by peritoneal macrophages and *Listeria monocytogenes* study). The combination of the failure of assays to produce reproducible results and the concern raised by the technical problems associated with conducting the assays represents a limitation in the conduct of the study.

3. All of the immunotoxicological data presented was evaluated by using an analysis of variance and Tukey's multiple comparison to compare control and treatment groups. Results from immunotoxicological assays do not always follow a normal distribution and, at times, a more appropriate evaluation would be one using a non-parametric analysis. Accordingly, this fixed approach used by the authors in evaluating their data is considered to be a weakness. A better procedure would be to use a decision tree approach. The data are first evaluated to determine if it is normally distributed and for homogeneity of variances using the Bartlett's Chi Square Test or another statistical test. Homogeneous data would be evaluated by a parametric one-way analysis of variance. When significant differences occur, treatment groups would be compared to the control using the Dunnett's t Test. Non-homogeneous data would be evaluated using a non-parametric analysis of variance. When significant differences occur, treatment groups would be compared to the control using the Wilcoxon Rank Test.

4. In general, there was sufficient data presented to confirm the findings in the report and usually the data could be tracked from the information provided in the tables to the final results in the figures. Areas of weakness include the lack of necessary information in several of the methods write-ups. More information needs to be provide on the CD4/CD8 surface marker analysis. Such as exactly which antibodies were used and an explanation why other cell types were not included in the evaluation. Why were B-cells and total T-cells not counted?. Furthermore, the surface marker data needs to also be presented as absolute numbers in addition to percent values.

More information needs to be provided in how the Lytic Unit was calculated for the natural killer cell assays. It appears that in some of the natural killer cell assays only three effector-to-target ratios were used. With such a small number of points defining the curve, accurate determination of Lytic Units is questionable as some of the values have to be extrapolated when 10% lysis of the target cells is used to define the Lytic Unit. Finally, it would be extremely helpful to have all the graphs for a particular assay drawn to the same scale. This would make interpretation across repeat studies and at the different time periods much easier for the reader.

5. The immunotoxicology data performed and reported to date are of sufficient quality for use in hazard characterizations. However, before a complete picture of the

effects of perchlorate exposure has on the immune system is determined, the results for the humoral immunity studies and additional host resistance studies need to be completed. When this data are obtained and added to the current base of information, we will have a significant understanding of the effects of perchlorate exposure and will have a better understanding of the potential for perchlorate to cause adverse effects on the immune system.

Since the immunotoxicological studies have only been conducted in young adult female mice, the data obtained will not provide information relevant to evaluating the sensitivities of specific subpopulations of exposed individuals. In general, the developing immune system of the fetus or newborn has been shown to be more susceptible to the effects of chemical and physical agents than young adult animals. None of the immunotoxicological studies conducted addressed the effect of perchlorate on the developing immune system of the fetus or the newborn.

6. Based on the written information provided to date, there still remains a significant amounts of critical data which needs to be obtained before an appropriate evaluation can be made of the effects of perchlorate exposure has on the immune system. Foremost among the critical data which must be evaluated are the humoral immunity results from the various plaque assays which need to be conducted. In addition, the host resistance studies which are to be conducted in the future will provide additional critical data in determining the effect perchlorate has on the immune response.

II. REVIEW OF TOXICOLOGICAL REVIEW DOCUMENT

A. EFFECTS OF CONCERN TO HUMAN HEALTH

1. The Toxicological Review document does a good job in adequately describing the key aspects of the protocol, as well as the conduct and the result of the various studies. Limitations were appropriately addressed and the discussion as written appears appropriate.

2. The analysis and evaluation performed on the immunotoxicological data were excellent and the EPA provided review of the data is considered to be one of the strengths of the Toxicological Review document. The only minor weakness is that concerns of the assay selected are not addressed in a very strong manner. Clarification on the "DTH" assay used in the immunotoxicology studies is addressed. There is a discussion regarding the need for humoral immune data in the evaluation of perchlorate; however, these comments could justifiably be made stronger or deleted if the humoral immune data are provided. Overall, it appears that the Toxicological Review document has covered the relevant studies and in the discussion appropriately addressed the various inconsistencies observed in the various studies.

3. Included in the immunotoxicology section of the Toxicological Review document was the additional statistical analysis conducted on the T4 levels obtained from the animals used in the study. In conducting the analysis the EPA personnel used a two-way ANOVA in evaluating the T4 data. The approach utilized and the interpretations of the results from the analysis appear appropriate and are consistent with effects observed on T4 in other species. The additional analysis contributes significantly to the overall data set of the effects of perchlorate on the hypothalamic-pituitary-thyroid axis.

4. The papers listed below are germane to the immunotoxicological approach used in selecting assays for evaluating the effects of test compounds, such perchlorate, on the immune system. They are submitted for consideration for enclosure in the Toxicological Review document.

Luster, M.I., Munson, A.E., Thomas, P.T., Holsapple, M.P., Fenters, J.D., White, K.L., Jr., Lauer, L.D., Germolec, D.R., Rosenthal, G.J. and Dean, J.H.: Development of a testing battery to assess chemical-induced immunotoxicity: National Toxicology Program's guidelines for immunotoxicity evaluation in mice. *Fund. Appl. Toxicol.*, 10:2-19, 1988.

Luster, M.I., Portier, C., Pait, D.G., White, K.L., Jr., Gennings, C., Munson, A.E. and Rosenthal, G.J.: Risk assessment in immunotoxicology. I. Sensitivity and predictability of immune tests. *Fund. Appl. Toxicol.*, 18:200-210, 1992.

Luster, M.I., Portier, C., Pait, D.G., Rosenthal, R.J., Germolec, D.R., Comment, C.E., Corsini, E., Blaylock, B.L., Pollock, P., Kouchi, Y., Craig, W., White, K.L., Jr. and Munson, A.E.: Risk assessment in immunotoxicology: II. Relationship between immune and host resistance tests. *Fund. Appl. Toxicol.*, 21:71-82, 1993.

C. ADDITIONAL ISSUES PERTAINING TO THE TOXICOLOGICAL REVIEW DOCUMENT

1. At the present time, the other sections of the document appear appropriate and I can not offer additional suggestions which would improve the quality of the document.

2. With the addition of the remaining data to be generated, the document is useful in characterizing the human health/ecotoxicological effects of ammonium perchlorate.

III. HAZARD CHARACTERIZATION

A. DEVELOPMENT OF REFERENCE DOSE (RfD)

1. Based on the data presented and the discussion included in the Toxicological Review document the individual NOAELs/LOAELs appear appropriate.

2. It seems prudent with the amount of data collected that the totality of the data on thyroid toxicity be considered for establishing the RfD and not base the results on the results of a single study.

3. Assuming that the outstanding data, plaque assay results and host resistance studies for example, only produce effects at higher dose levels, then the principle study selected by the EPA appears appropriate. However, additional documentation is needed to justify the manner in which the data were handled.

4. Based on the information provided and the discussion put forth in the Toxicological Review document, a factor of 3 seems appropriate.

5. It is unclear from the Toxicological Review document why a factor of 3 and not 10 is used for the intrahuman variation. If data are available on this point, it needs to be included in the discussion or the discussion modified so that the reasoning is more clear to the reader.

6. As presented in the document, the data support this position.

IV. FURTHER TESTING NEEDS FOR PERCHLORATE

A. TOXICOLOGICAL TESTING

1 The experimental designs appear adequate to identify the potential hormone disturbing effects on development and reproductive performance.

2. If in the future immunotoxicological studies to be conducted, the humoral immune response, as evaluated in the plaque assay, is found to be a target for perchlorate, then an appropriate host resistance model which is defended by antibody should be considered. Among the assays which may be considered to meet this objective is the *Trichinella spiralis* model, *Streptococcus* model, and the rodent malaria models.

In the immunotoxicological studies presented, the "phagocytic" ability of the peritoneal macrophages was identified as a compromised function following exposure to perchlorate. However, unlike aveolar macrophages, in which decreased phagocytosis correlates with decreased host defense to lung pathogens, there is not a good correlation between decreased peritoneal macrophage phagocytosis and decreased host resistance to systemic pathogens. To confirm that the exposure to perchlorate does adversely effect this macrophage function, evaluation of the phagocytic ability of the fixed tissue macrophages could be considered in the *in vivo* RES assay which measure the functional ability of the reticuloendothelial system.

3. The development of such a model would have the potential to help establish more appropriate uncertainty factors which could be used in establishing acceptable levels of perchlorate in the environment.

Statistical Analysis Issues in Perchlorate Studies

Prepared by Joe Haseman, Ph.D.
National Institute of Environmental Health Sciences
Biostatistics Branch
Research Triangle Park, NC 27709

GENERAL COMMENTS

(1) Because there were literally hundreds of pages of data, graphs, tables, and statistical analyses, and limited time to evaluate them, I elected to concentrate on the data most relevant to the determination of the revised RfD, namely, thyroid histology observed in pups on PND5 in the neurodevelopmental study. The EPA statistical analyses of T3, T4, and TSH in various studies and the motor activity data were also evaluated.

(2) A very important preliminary issue, but one beyond my area of expertise, is the selection of the response variable upon which to base an RfD. Is an apparently transient effect on thyroid histology (cell hypertrophy, lumen size) seen on PND5 really the appropriate variable to be used to estimate a revised RfD?

(3) Assuming the answer is Yes, then I am uneasy regarding the revised RfD of 0.0009 mg/kg-day for perchlorate. My concern is that it may be too high, since statistically significant ($p<0.05$) effects were observed for thyroid follicular cell hypertrophy and lumen size at the lowest dose evaluated (0.1 mg/kg-day), despite sample sizes as low as 6 animals per sex per dosed group. One can only speculate as to what might have been observed at even lower doses and/or with more animals per group. In fact, significant reductions in T3 and T4 in the 90 day study were seen at doses ten times lower (0.01 mg/kg-day), with no NOAEL being found. I would feel much more comfortable had a NOAEL been established and the "composite uncertainty factor of 100" applied to a NOAEL rather than to a LOAEL.

(4) There were quite a few problems with the statistical analysis and data reporting. In general, these problems resulted in an understated statistical significance for the various effects evaluated. Certain figures in the EPA Review Draft also require revision. These matters are discussed in more detail below.

SPECIFIC COMMENTS

A. Thyroid epithelial cell hypertrophy ("standard histology")

The package we were given included the EPA statistical analysis of this variable (Marcus, 1998) and the summarization of the test results in the EPA Review Draft. For each of five dose levels (including controls) six males and six females were classified into one of three hypertrophy categories (Level 0, 1, or 2) and contingency table analyses were carried out. My concerns with this analysis are summarized below.

(1) In Marcus Table 1 a female is missing from the 10 mg/kg-day dose group. The correct pattern of response in this group is 0-5-1, not 0-4-1 for Levels 0, 1, and 2 respectively.

(2) The p value reported by Marcus for the Table 1 analysis of females ($p=0.026$) is also incorrect; this is the p value for the males (analyzed in Table 2).

(3) Similarly, the p value reported by Marcus for Table 2 is incorrect for two reasons: (a) it reflects the Table 1 data, not the Table 2 data, but (b) because of the raw data error noted in (1) above, it is incorrect in any event. It should be p=0.142, not p=0.218.
(4) Marcus Table 3 (both sexes combined) is also incorrect, since (a) there is an extra animal in the 0.1 mg/kg-day group; the correct pattern of response is 4-6-2, not 4-7-2; (b) there is a missing animal in the 10 mg/kg-day group as noted previously; the correct pattern of response is 0-7-5, not 0-6-5, and (c) because of these errors the p value is incorrect and should be p=0.010.
(5) The EPA Draft Report further compounds the problem by incorrectly reporting the results of these analyses, stating, for example (page 5-28), that "when data on both sexes were combined, the lowest dose, 0.1 mg/kg-day was significant at 0.012 (df=8)." This p value does not reflect an effect at 0.1 mg/kg-day, but instead reflects the significance of the difference in the overall pattern of response among the five groups. Subsequent pairwise comparisons carried out by Marcus reflect the individual dose comparisons.
(6) Unfortunately, these subsequent pairwise comparisons needlessly collapse the data over response categories and thus are relatively insensitive for detecting treatment effects. For example, when comparing the pattern of response in the lowest dose (0.1 mg/kg-day) and control male groups, the two patterns of response (5-1-0 vs 1-4-1 for Levels 0, 1, and 2 respectively), are collapsed to 5-1 vs. 1-5. Thus, the Marcus analysis considers the Level 2 response observed in one low dose male as being the same as a Level 1 response. However, there is no reason not to preserve the complete ranking of the data. Marcus reports a (two-sided) p value of p=0.080, whereas a test that maintains the ordering of the responses (Cochran-Armitage trend test) gives a corresponding p value of p=0.024). The statistical significance of the combined (male + female) analysis is similarly underreported.
(7) Because this variable (and lumen size discussed below) are critical to the selection of the RfD, the EPA Draft Report should bring forward and include in the document some of the critical (corrected) contingency tables from the Marcus report. In my opinion, the current table that summarizes these data (Table 5-3; page 5-29) provides insufficient detail to critically evaluate variables that are so important. Moreover, no statistical results are indicated in that table.
Interestingly, despite all of these problems, the correct decision was reached, in my opinion, for this variable, namely, that no NOAEL has been established and that the lowest dose, 0.1, was a LOAEL (page 5-29 of EPA Draft Report). I agree with this conclusion, and it is especially important, since this low dose effect was seen despite the relatively small sample sizes of 6 animals per group (or 12 per group if males and females are pooled). Perchlorate was clearly affecting thyroid epithelial cell hypertrophy at doses as low as 0.1 mg/kg-day.
B. Lumen size ("standard histology")
The package we were given included the EPA statistical analysis of this variable (Marcus, 1998) and the summarization of the test results in the EPA Review Draft. For each of five dose levels (including controls) six males and six females were classified into one of four categories (Level 0, 1, 2, or 3) and contingency table analyses were carried out. My concerns with this analysis are summarized below.

(1) The Likelihood ratio Chi-square p value for females (Marcus Table 4) should be p=0.416, not p=0.218. The corresponding p value for males (Marcus Table 5) should be p=0.011, not p=0.007. P=0.007 is the result of the "usual" Chi-square test, but Marcus explicitly states the results he reports are for the likelihood ratio [Chi-square] test, and these are indeed the p values he reports for the hypertrophy data discussed above.
(2) The combined data severities shown in Table 5-3 (page 5-29) of the EPA Draft Report have three minor roundoff errors: 0.084 should be 0.083, 0.066 should be 0.067, and 2.16 should be 2.17.
(3) There are two errors in the combined data table (Marcus Table 6). The correct pattern of response for the 3 mg/kg-day group is 1-3-6-2, not 3-3-6-2, while the correct response pattern in the 10 mg/kg-day group should be 0-2-6-4, not 0-3-6-4. These errors are apparently just typos, since the overall p value (p=0.008) appears to be correct.
(4) The EPA Draft Report reports certain of these results incorrectly, stating, for example (page 5-28), that "when data on both sexes were combined, the lowest dose, 0.1 mg/kg-day was significant .... at 0.008 (df=12)." However, this 12 df Chi-square p value simply reflects the significance of the difference in the overall pattern of response among the five groups. Subsequent analyses carried out by Marcus reflect the individual dose comparisons.
(5) These subsequent statistical analyses are relatively insensitive, since they needlessly collapse the data over response categories. For example, when comparing the pattern of response in the pooled lowest dose (0.1 mg/kg-day) and control groups, Marcus takes the pattern of response in the two groups (6-4-2 vs 2-6-4 for Levels 0, 1 and 2 respectively), and collapses them to (6-6 vs. 2-10), which he incorrectly reports in his Table 9 as 6-2 vs. 2-10. However, there is no reason not to preserve the complete ranking of the data. Marcus reports a (two-sided) p value of p=0.193, whereas a procedure that maintains the ordering (Cochran-Armitage trend test) gives a corresponding p value of p=0.106. Note that, contrary to the statement made in the EPA Draft Report (page 5-28), this effect at the 0.1 mg/kg-day is not statistically significant at the $p < 0.05$ level.
Interestingly, despite all of these problems (and the lack of statistical significance at the lowest dose), the EPA made the correct decision, in my opinion, for this variable, namely, that no NOAEL has been established and that the lowest dose, 0.1 mg/kg-day, was a LOAEL (page 5-29). Despite the lack of statistical significance at the lowest dose, the morphometric analysis of lumen size (discussed below and based on additional animals) did in fact show a statistically significant effect at this dose (0.1 mg/kg-day).
C. Lumen Size (morphometric analysis)
(1) The figure in the EPA Review Draft that presents these data (Figure 5-10, page 5-31) is misleading in that the legend indicates that the data are being presented for both females (circles) and males (squares). The actual graph contains only circles, but further investigation of the raw data reveals that this graph does not correspond to females but rather to the pool of males and females. The legend should be deleted and the title of the figure modified to make it clear that the data are for pooled males and females. This was a common error that occurred in many of the figures in the EPA Draft report.
(2) The ANOVA result for "treatment" presented in Figure 5-10 (page 5-31) is the wrong ANOVA upon which to base a formal statistical analysis of dose effects. There are three primary "main effects" for these data:

(a) sex (males vs. females). Males and females were not significantly different nor were the interactions involving sex significant. Thus, it is reasonable to collapse the data over sexes as the EPA has done in Figure 5-10.
(b) "block" (Block 1 vs. Block 2). These data were generated in two separate blocks or replicates, the first containing six animals per group, the second containing four animals per group. These two blocks were significantly different ($p<0.01$), so block must be retained as a significant source of variability and not pooled with error.
(c) dose. Clearly, the five dosed groups are significantly different ($p<0.001$).
The EPA statistical analysis pooled the block sum of squares (and all interactions involving block) with error, thereby artificially inflating the error term and resulting in an insensitive test that understated the true significance of the treatment effect. This was an important recurring problem with many of the EPA statistical analyses. If a significant main effect has been identified (e.g., gender, block, or day), then it is inappropriate to pool the corresponding sums of squares with error when evaluating other main effects (e.g., treatment) by ANOVA.
(3) The pairwise comparison results presented in Figure 5-10 (page 5-31) are incorrect in that they do not adjust for differences between blocks and understate the dose effects. When the variability between blocks is taken into account, the low dose (0.1 mg/kg-day) shows a significant ($p <0.05$) reduction in lumen size. The reduction at the 1.0 mg/kg-day dose is not quite significant. Thus, I do not agree with the EPA conclusion that "the NOAEL for thyroid histopathology, based on the morphometric assessments, is 1.0 mg/kg-day" (page 5-30). In my opinion no NOAEL has been established and 0.1 mg/kg-day is in fact a LOAEL.
(4) Interestingly, the dosed group effects on lumen size are stronger in the second block (4 animals per group per sex) than in the first block (6 animals per group per sex). Since the first block apparently served as the basis for the statistical analyses described in A. and B. above, had "standard histology" been carried out for the second block, it is likely that the low dose effect on lumen size described in B. above would have been statistically significant.
(5) It is not clear what multiple comparison procedure is being used for these data. The original source of these data provided to us is Crofton (1998f) [not Crofton (1998e) as stated in page 5-30 of the EPA Review Draft]. Crofton (1998f; page 2 of 29) reports that the data were analyzed by "Turkey's Studentized Range Test." This should be "Tukey" rather than "Turkey", but the actual statistical analysis carried out and reported by Crofton (1998f) for these data is not Tukey's Studentized Range test but rather Duncan's Multiple Range test (see pages 23, 26, and 29 of Crofton, 1998f). Either of these multiple comparison procedures would be OK, but they are different tests.
(6) There are additional lumen size data from PND 90-92 that were reported as non significant by Channel (1998b) but have not yet re-assessed by the EPA. My reanalysis of these data suggests that the conclusion of Channel (1998b) is correct. That is, although there is a highly significant sex effect (females > males), there does not appear to be a dose effect. Thus, I agree with the EPA conclusion (page 5-27) that "these data suggest a recovery from the effects observed in the thyroids of pups at PND5."

D. Thyroid and Pituitary Hormone Data: T3

(1) The figure in the EPA Review Draft that presents these data (Figure 5-11, page 5-33) is misleading in that the legend indicates that the data are being presented for both females (circles) and males (squares). The actual graph contains only circles, but this is not the female data. In fact, according to page 5-31 of the EPA Review Draft "male and female samples were pooled for each litter", so presumably it is not possible to compare male and female responses for T3 as is implied in Figure 5-11.

(2) According to page 16 of 19 in Crofton (1998e), which is the source of the T3 data analysis, the overall F for treatment effects is 216.89, not 214.89 as indicated in Figure 5-11.

(3) The EPA uses a rather conservative procedure to assess dose effects. Not only do they use a multiple comparison procedure to adjust for the number of doses (Tukey's test), they further adjust the p value down to p=0.028 by dividing "the alpha of 0.05" by the square root of "the number of ANOVA tests carried out " (three). This ad hoc adjustment requires a reference. The Bonferonni adjustment divides alpha by the number of tests carried out, but I am unaware of a multiple comparison procedure that divides alpha by the square root of the number of tests carried out. Such an adjustment is unnecessary in my opinion and was apparently not used in other EPA analysis (e.g., thyroid hypertrophy and lumen size). In fact, I cannot confirm that it was used in these analyses, since the SAS output in Crofton (1998e) apparently uses an alpha of 0.05 not 0.028 for Tukey's test. Similar comments apply in other instances in which this alpha adjustment was apparently used. This adjustment makes no difference for T3, but does for T4. I agree with the EPA conclusion that the NOAEL for T3 is 0.1 mg/kg-day.

E. Thyroid and Pituitary Hormone Data: T4

(1) The figure in the EPA Review Draft that presents these data (Figure 5-12, page 5-34) is misleading in that the legend indicates that the data are being presented for both females (circles) and males (squares). The actual graph contains only circles, but this is not the female data. In fact, according to page 5-31 of the EPA Review Draft "male and female samples were pooled for each litter", so presumably it is not possible to compare male and female responses for T4 as is implied in Figure 5-12.

(2) The conservative multiple comparisons procedure used by the EPA is discussed above. An alternative procedure (Protected Fisher's LSD) would identify the 1.0 mg/kg-day dose as producing a significant effect, even using the p=0.028 "correction factor", so I disagree with the EPA conclusion (page 5-32) that this dose is a NOAEL for T4. In my opinion the NOAEL is the same for both T3 and T4, namely, 0.1 mg/kg-day.

F. Thyroid and Pituitary Hormone Data: TSH

(1) The figure in the EPA Review Draft that presents these data (Figure 5-13, page 5-35) is misleading in that the legend indicates that the data are being presented for both females (circles) and males (squares). The actual graph contains only circles, but this is not the female data. In fact, according to page 5-31 of the EPA Review Draft "male and female samples were pooled for each litter", so presumably it is not possible to compare male and female responses for TSH as is implied in Figure 5-13.

(2) I agree with the EPA conclusion that the only statistically significant effect is at the top (10 mg/kg-day) dose (Figure 5-13; page 5-35).

G. Motor Activity Data

(1) In my review of the raw data, I found a minor error: On page 611 of the Argus 1613-002 submission, the total number of movements for control female 1750 is 108, not 230, as reported in Table F8. This error had little impact on the subsequent statistical analysis.
(2) The figure in the EPA Review Draft that presents these data (Figure 5-14, page 5-36) is misleading in several ways. First, it has two legends. The first legend indicates that the data are being presented for both females (circles) and males (squares). The second legend indicates that the data are being presented for both movements (circles) and time (triangles). Are the circles presented in this figure for females or for movement? The second legend appears to be the correct one, and the first legend should be deleted. Moreover, the figure should state explicitly that these data are for males only. Importantly, the female data (not shown) show the opposite trend from that of the males. To be completely fair, the figure should show the female response also, or at the very least, it should be mentioned in the text.
(3) Since PND14 is the only time showing a possible motor activity effect (there were no changes detected at PND18, PND22, or PND59; see page 5-32 of EPA Draft Report), and this nonsignificant increase in high dose males is not seen is in females, I strongly doubt that this slight increase is a true perchlorate effect.
(4) The EPA Review Draft expresses surprise that the high dose motor activity effect in males is not significant and requested from Argus Laboratory and their sponsor an additional statistical analysis "using gender as a within subject variable, or alternatively, using a nested design with gender nested under litter" (page 5-33). Argus Laboratory responded (York, 1998b) by carrying out an additional statistical analysis using gender as a between-subject variable. Since the females show the opposite effect of males, the overall dose effect is not significant, as one would expect (the male effect considered individually is also not significant). The EPA expressed disappointment, noting that the "secondary analysis submitted is still not what the EPA requested" (page 5-34).
The EPA (page 5-33) cites two references (not provided) to support their requested analysis, but it makes no sense to this reviewer to attempt to treat gender as a nested factor. A classic example of nesting is a reproduction or teratology study in which litters are nested within treatments and fetuses are nested within litters. This is because a treatment may have many unique litters within it, and a litter may have many unique fetuses within it. However, a litter does not have many unique sexes within it. Gender is a fixed, not a random effect. and should be considered a crossed, not a nested factor. In my view the Argus analysis is correct, and there is no significant dose effect.
(5) The EPA also faults Argus Laboratory and their sponsor for failing to respond adequately to the request for an explanation of why the statistical analysis failed to detect significance in the PND14 motor activity for male rats (page 5-34). The primary reason for the lack of statistical significance is that these data show considerable animal-to-animal variability in motor activity, with the SD actually exceeding the mean in certain groups. The EPA Draft report notes this large within group variability (page 5-37), but still concludes that the "increase in activity should be considered biologically significant until additional data can be marshaled to suggest or prove otherwise" (page 5-38). Ultimately, the decision of whether or not a non-significant increase is biologically important is a matter of scientific judgement, but I feel that the alternative explanation given on page 5-37 that this increase "may indeed be a Type I error and would not be

found again if this experiment was repeated" is more nearly correct. However, referring to this result as a "Type I error" implies a statistically significant effect, whereas none exist for these data.

(6) Another factor that supports my point of view on this matter is that it is the male controls that appear to be unusually low, not the high dose group that is unusually high. For example, the mean motor activity response in the high dose male group (Number of movements = 404; Time spent in movement=364) is very similar to the mean control female response (Number of movements = 393; Time spent in movement = 430; see Argus, Appendix F, pages 509-512). There is nothing in these data that would suggest a gender difference in motor activity. In summary, in my opinion this non-significant increase in motor activity in one dosed group in one sex at one time period simply reflects random variability, nothing more.

(7) I might mildly fault one aspect of the Argus analysis: the data show evidence of heterogeneity of variance across groups, which should have been dealt with by a data transformation, such as a logarithmic or square root transformation. However, such a transformation would not have affected the negative result obtained.

In summary, I consider the motor activity data to be totally negative.

H. 90 Day Experiment: Thyroid and Pituitary Hormone Analyses

The EPA Draft Report (page 5-17) states that "the Crofton (1998b) analysis contains a printout of all the individual animal data", but that printout was not included in the information provided to me; only summary statistics and test results are given. This was especially frustrating since the summary statistics given by the EPA differed in some cases from the summary statistics for the same data reported by the test laboratory (Springborn). Since there were several problems with the EPA statistical analysis, I have done the best I can with the summary statistics provided and when discrepancies existed, I assumed the EPA data were correct and the Springborn data were incorrect.

The EPA reportedly used the same conservative multiple comparisons adjustment procedure described in D (3) above, including the "correction" of dividing the alpha of 0.05 by the square root of three (page 5-17). However, the SAS output contained in Crofton (1998b) for Tukey's test gave no indication that this adjustment was used.

H-1 T3

(1) In Figure 5-19 for Day 14 (page 5-23), the plotted points for the 0.05 and 0.2 males have been interchanged (see page 50 of 95 of Crofton, 1998b). I note that this is one instance in which the mean responses reported by the EPA (138.0 and 141.2) differ from the values reported by the test laboratory for these same two groups (140 and 139; see page 447 of SLI Study No. 3455.1).

(2) In Figure 5-19 (page 5-23) the female data do not contain letters showing which groups are statistically different (in this case, there is no significant difference among dosed groups). Moreover, the legend to the Day 120 plot is incorrect and shows males and females separately, whereas the plot is for males and females combined.

I agree with the EPA conclusion (page 5-22) that no NOAEL for T3 has been established for males and females at Day 90 and for males at Day 14. Importantly, despite the

conservative statistical methodology used, the lowest dose tested (0.01 mg/kg-day) still shows significant reductions in T3 (Figure 5-6; page 5-19).

H-2 T4

(1) I do not agree with the EPA statistical analysis of this variable. Figure 5-7 (page 5-20) states that "the data were collasped [sic] across gender", because of a highly significant day by treatment interaction. This makes absolutely no sense. This variable showed both a highly significant ($p<0.0001$) gender effect and a highly significant ($p<0.0001$) gender by treatment interaction. Thus, it is critical that each sex be examined separately. Including gender sums of squares with the error term results in a very insensitive (and misleading) test. Figure 5-7 should be modified accordingly.

(2) Based on the Springborn summary statistics, it appears that the EPA is correct that the significant reduction in T4 at Day 14 is limited to the top dose group. What Figure 5-7 (page 5-20) does not show is that the response in males is consistently higher (at all doses) than the response in females.

(3) Based on the Springborn summary statistics, it appears that the EPA is correct that on Day 90 all doses tested, even the low dose of 0.01 mg/kg-day, significantly reduced T4. What Figure 5-7 does not show is that this reduction is significant in both males and females and that the T4 values for males are generally higher than for females.

(4) Based on the Springborn summary statistics, it appears that the EPA conclusions that on Day 120 T4 is significantly reduced at all dose levels is true for males, but not for females (this is the strongest evidence for the gender by treatment interaction found by the overall test). For females, the reduction in T4 appears to be limited to the top dose.

(5) Figure 5-8 (page 5-21) states that because "there was a significant gender*treatment interaction, data was collapsed across days." This makes no sense. The day effect and treatment by day interaction are both highly significant ($p<0.0001$), so as noted previously, it is wrong to include these effects as part of the error term. This problem is compounded by the fact that (i) not every dose is represented on each day, so comparing treatment means from one day with control means on other days is misleading when there are highly significant day effects, and (2) the significant ($p<0.0001$) treatment by day interaction implies that the treatment effect varies from day to day.

In summary, comparisons should be carried out separately for each time period, and the results in Figure 5-8 are misleading. What is needed is a figure like Figure 5-7, but with males and females presented and evaluated separately.

Despite these problems, I agree with the EPA conclusion (page 5-22) that "The LOAEL, based on decreases in T3 and T4 at 90 days, is 0.01 mg/kg-day. No NOAEL could be calculated for T3 and T4." If these are scientifically important variables relevant to setting the RfD, then I would be very uncomfortable with the recommended RfD of 0.0009 mg/kg-day, a level that is only 11-fold lower than a dose demonstrated to have a statistically significant adverse effect. Is this margin of safety adequate?

H-3 TSH

(1) I do not agree with the EPA conclusion (page 5-22) that "The NOAEL for TSH is 0.05 mg/kg-day." Not only was TSH numerically elevated at all three time points in both sexes at 0.01 and 0.05 mg/kg-day, it was significantly ($p<0.05$) elevated at the 0.05 mg/kg-day dose for females at 14 days (see Figure 5-9; page 5-23) and in females at 120 days.

Clearly, (in my view) 0.05 mg/kg-day is not the NOAEL for TSH, and probably not even a LOAEL (see final point discussed below).
(2) In the first panel of Figure 5-9, Day 15 should probably be Day 14 (consistent with the figures for T3 and T4). The test results appear to be OK as reported.
(3) In the second panel of Figure 5-9, the male/female legend should be deleted, since the data are plotted for combined males and females. In this instance, it is acceptable to combine males and females, since neither the gender effect nor the gender x treatment interaction was significant.
(4) In contrast, in the third panel of Figure 5-9 page 5-23), the gender effect is highly significant ($p<0.0001$), so it is wrong to ignore this source of variability and include it as part of the error term. Each sex should be considered separately, and the elevated TSH is statistically significant in all three dosed groups for females. For males, the elevated values are not statistically significant (see Springborn, 1998). The EPA correctly notes (page 5-17) that "the new EPA analysis failed to detect a significant effect of perchlorate on TSH at the 120-day time point," but this was because the EPA analysis was flawed as noted above. The Springborn analysis is more nearly correct in this instance. Also, in this third panel (Figure 5-9; page 5-23) the male/female legend should be deleted, since the data are plotted for males and females combined.
(5) As a concrete example of the misleading results obtained by pooling the gender effect with error, on page 41 of 95 of Crofton (1998b), the ANOVA indicates that the treatment effect on Day 120 for TSH (after adjusting for gender differences) is highly significant: $F(3,71)=7.63$, $p=0.0002$. The gender difference is even more significant $F(1,71) = 310.13$, $p<0.0001$. This EPA analysis is appropriate and indicates a highly significant treatment effect. However, the EPA then incorrectly pools the huge gender sum of squares (and the gender by treatment sum of squares) with error and reanalyzes the data. The resulting error term increases more than 5-fold, and the treatment effect is now no longer significant: $F(3,75)= 1.50$, $p=0.2224$; see page 79 of 95 of Crofton (1998b) and Figure 5.9, page 5-23 of the EPA Draft Report. Thus, the EPA concludes that the treatment effect it found and reported earlier for these data is not significant after all.
However, the treatment effect is real, as indicated by the first ANOVA result given above. The EPA analysis cannot detect it because the gender effect is so strong, that regarding it as error greatly diminishes study sensitivity.
(6) Since the raw data were not included (despite the contrary remark on page 5-17), I could not ascertain with certainty whether or not the consistently elevated TSH's in the 0.01 mg/kg-day group are significant or not (considered globally over sex and day). My best guess is that these elevated values would be marginally significant and thus (as for T3 and T4) no NOAEL has been established for TSH for these data.
I. 14 Day Experiment: Thyroid and Pituitary Hormone Analyses
The variables analyzed in Figures 5.1 through 5.5 of the EPA Draft Report are T3, T4, TSH, rT3, and hTg respectively. Certain of the same problems existed in these data that existed in the other statistical analyses discussed above.

I-1 T3

I agree with the EPA conclusion (page 5-11) that no NOAEL was established for this variable and that the lowest dosage of 0.1 mg/kg-day was a LOAEL. It is interesting that at this low dose, the reduction in T3 is quite striking in females but not statistically significant in males (Figure 5.1; page 5-10).

I-2 T4

I do not agree with the EPA decision to pool males and females despite a highly significant ($p<0.0001$) gender effect. This matter has been discussed previously and will not be repeated here. Figure 5-2 (page 5-11) should present results for males and females separately, not pooled. The current Figure 5-2 is for pooled males and females and thus the legend indicating separate plots for males and females should be deleted from the figure. Despite the insensitivity of the statistical analysis, the reduction in T4 at the lowest dose tested, 0.1 mg/kg-day, was still significant (as for T3), and I agree with the EPA conclusion (page 5-11) that no NOAEL has been established for T4.

I-3 TSH

This analysis (Figure 5-3; page 5-12) seems basically OK. Arguably, 0.1 mg/kg-day is a NOAEL for TSH.

I-4 rT3

Since males and females showed no significant difference in response (nor was there a gender* treatment interaction) pooling sexes is appropriate for this variable. The legend in Figure 5-4 indicating plots for both sexes should be deleted. The plot is for males and females pooled and shows that 0.4 is arguably a NOAEL for this variable. I do not understand how it is possible for the EPA to conclude that 0.17 was the NOAEL (see Table 5-2; page 5-14), when the lowest doses actually used for males/females were 0.11/0.12 and 0.44/0.47. The EPA Draft Report does not explain how the NOAEL was (apparently) interpolated between these two doses. I was under the impression that a NOAEL had to be one of the doses actually used in a study, not an interpolated or extrapolated value.

I-5 hTg

I agree with the EPA conclusion (see Figure 5-5; page 5-14) that no NOAEL was established for this variable and that the lowest dosage of 0.1 mg/kg-day was a LOAEL.

J. T3, T4, and TSH in pregnant New Zealand Rabbits.

(1) The EPA Draft Report states (page 5-45) that these data were evaluated by Tukey's Studentized Range test. However, the statistical analyses of these data that are enclosed and cited (Crofton, 1998h) report the results of Duncan's Multiple Range Test, not Tukey's test.

(2) The EPA Draft Report states (page 5-45) that a multiple comparisons correction was used (alpha=0.0289, not alpha=0.05), but the Duncan Multiple range test was carried out using alpha=0.05, not alpha=0.0289. If such a "correction" for multiple comparisons is to be used (and I do not feel it is necessary, as noted previously), then it should be applied to the individual pairwise comparisons, not just to the overall test.

(3) There are main effects (e.g., dose and gender) and there are interactions (e.g., dose by gender), but it is unclear what the EPA means by an "interaction main effect" (page 5-45). In any case, these data had only one main effect (dose), so no interactions are possible.

Thus, I do not understand what "interaction" the EPA is referring to for these data. This unusual terminology ("interaction main effect" is used throughout the EPA Draft Report.
(4) The T3 data do indeed appear to be negative, as shown in Figure 5-15 (page 5-46). However, since these data are for pregnant rabbits, the legend stating that male data are also being graphed should be deleted.
(5) The T4 (Figure 5-16, page 5-47) and TSH (Figure 5-17, page 5-48) data appear to be analyzed correctly, but the legend in both figures should be deleted. It is interesting that T4 shows significant reductions at four different dose levels, while T3 shows no significant effects at all.

K. 14- and 90-Day Exposures to Ammonium Perchlorate in B6C3F1 mice: T4 and TSH
These data (Figure 5-18, page 5-52 and Figure 5-19, page 5-53) appear to be evaluated satisfactorily. There apparently were no T3 data.

CONCLUSION

I am uneasy regarding the selection of an RfD for perchlorate. Specifically, I am concerned that it may be too high. Recall that
(1) The lowest dose tested 0.1 mg/kg-day produced significant effects on T3, T4, and hTg in the 14-day study.
(2) A ten-fold lower dose (0.01 mg/kg-day; the lowest dose tested) produced significant changes in T3, T4, and possibly TSH in the 90 day study.
(3) On PND5 the lowest dose tested (0.1 mg/kg-day) significantly affected both thyroid follicular epithelial cell hypertrophy and lumen size. These appear to be the variables upon which the RfD was based.
As noted earlier, I do not favor a policy that applies a 100-fold safety factor to a dose that is still producing significant changes in key variables at the lowest dose tested with as few as six animals per group and with no NOAEL established.
Finally, as noted in my General Comments, a critical issue beyond my area of expertise is the selection of the endpoint upon which to base the RfD. If the endpoint were some obvious adverse health effect such as cancer or malformations, then there would be less of a problem. But, thyroid hypertrophy and lumen size? Playing devil's advocate, I raise the following question: Is a (possibly) transient (the EPA Draft Report notes on page 5-27 that "these data suggest a recovery from the effects in the thyroids of pups at PND5") effect on thyroid histology the proper variable upon which to base an RfD? The answer may well be Yes, but this issue certainly needs to be considered carefully.

Joe Haseman
Mathematical Statistician
NIEHS
January 29, 1999

SUMMARY OF MAJOR STATISTICAL QUESTIONS/ISSUES

(1) In their comparison of individual dosed groups vs. controls for thyroid follicular cell hypertrophy and lumen size, why did the EPA collapse the data into 2 x 2 tables rather than preserve the complete ranking of the response variable? This collapsing of the data sacrificed information and reduced study sensitivity for detecting dosed group effects.

(2) In many of the EPA multi-factor ANOVA's, highly significant ($p < 0.0001$) gender, day, or block effects were found. In many of these cases, the EPA then ignored these highly significant sources of variability, and pooled these sums of squares with the error term when evaluating differences among dosed groups. This had the effect of inflating the error term and reducing study sensitivity for detecting dosed group effects. I do not understand how this approach to data analysis can be justified.

(3) What is the specific reference to justify the EPA's multiple comparison adjustment of dividing alpha by the square root of the number of variables evaluated? This is not the Bonferonni adjustment, which divides alpha by the number of variables evaluated. Also, why was the EPA selective in the use of this correction, using it in some data evaluations, but not in others? Finally, exactly how was this correction used? The EPA Draft report states that it was applied to all "interaction main effects", but it is unclear what this means. It was apparently not applied to Tukey's test, but if such a correction is to be used at all, it should apply to all pairwise comparisons as well as to tests for overall effects.

(4) Did the EPA carry out any preliminary heterogeneity tests prior to their ANOVA's? I could find no mention of such tests, and for certain of the variables it appeared that a data transformation was needed to equalize the variances.

(5) How can the selection of a NOAEL be justified for situations in which the selected NOAEL actually significantly affected the variable of interest (e.g., a NOAEL of 0.05 mg/kg-day for TSH in the 90 day study; see page 5-22, despite significant effects at this dose shown in Figure 5-9; page 5-23)?

(6) For the motor activity data, how does the EPA justify treating sex as a factor nested within litters rather than as a crossed factor? Specifically, what statistical analysis does the EPA propose for these data?

(7) There are some issues of consistency in the choice of statistical methodology. For example, was Tukey's Studentized Range Test used throughout (despite the apparent use of Duncan's Multiple Range Test in some instances)? Also, was the likelihood ratio (Chi-square) test used for all contingency table analyses, despite the apparent use of the "usual" Chi-square test in some cases? Also, how did the EPA in at least one case apparently interpolate the NOAEL (e.g., Table 5-2; page 5-14) rather than select one of the doses actually used in the study?

PEER REVIEW:

ECOLOGICAL RISK ISSUES: EPA DECEMBER 31, 1998 DRAFT: PERCHLORATE ENVIRONMENTAL CONTAMINATION: TOXICOLOGICAL REVIEW AND RISK CHARACTERIZATION BASED ON EMERGING INFORMATION

Prepared By
Rick Cardwell, Ph.D.
Parametrix, Inc.
Kirkland, Washington
January 26, 1999

ECOTOXICOLOGICAL EFFECTS OF CONCERN (MANUSCRIPT BY SPRENGER ET AL. (1998)

1. **Have the Key Aspects of Protocols, Methods and Ecotox Study Results Been Described Adequately?**

Response: EA Engineering, Science and Technology, Inc.: EA conducted the following tests: fathead minnow (*Pimephales*) and *Daphnia magna* acutes, fathead minnow 7-day chronic, *Ceriodaphnia* 7-day acute and chronic and 28-day lettuce chronic. The test protocols and methods used by EA Engineering, Science and Technology have been adequately described; more importantly, they conform to standard methods (e.g., ASTM and EPA). The test results meet the protocol requirements and are of good quality. The range-finding and definitive test results are in reasonable agreement, lending further confidence to the validity of the definitive test results.

Block Environmental Services, Inc.: BEA conducted 6-7 day chronics with fathead minnow and *Ceriodaphnia* in essentially the same manner as EA. The protocol and test documentation were significantly less extensive than EA. No protocol was provided, although BEA said it was available and similar to methods described in a EPA chronic test method manual. Only data summaries were provided, and the results could not be independently verified. However, the *Ceriodaphnia* test results were very close to those reported by EA, and the *Pimephales* data were within a range (factor of 3-4) that is usual for inter-laboratory variability. This study is believed to be of acceptable quality.

Dumont and Bantle (1998): Drs. Dumont and Bantle studied mortality and malformations in the African toad *Xenopus* exposed to aqueous perchlorate. Although a formal test protocol did not accompany the report, it was performed according to ASTM and contained extensive test data. These data appear to be of acceptable quality.

Nzengung (1998): Dr. Nzengung studied the uptake and metabolism of perchlorate in three woody plants (willow – *Salix*; Eastern cottonwood – *Poplar*; and *Eucalyptus*), French Tarragon, spinach, the aquatic plant *Myriophyllum*, and microbial mats. There appeared to be no formal protocol, but the study was described adequately. These results should be treated as qualitative and indicative of trends, not as absolute values, because perchlorate concentrations in the medium were not constant. Rather, they declined with time as the initial addition of perchlorate was not renewed during the course of the studies, and evidence was provided it was taken up by the plants as well as biodegraded.

2. **Have Limitations in Studies Been Appropriately Discussed?** There appeared to be two salient limitations in these tests, and they concern the chronic tests of *Ceriodaphnia* and *Pimephales* and Nzegung's (1998) phytoremediation study. I do not consider the 7-day chronics as definitive because their exposure durations were short, and for one species (*Pimephales*), they did not encompass the species' life cycle. With a chemical that affects hormonal function, I would have liked to have seen a multi-generation test with *Ceriodaphnia* and a full life cycle test with *Pimephales*. The question: were the 7-day test durations sufficient to check on hormonal effects with fish and invertebrates that had no prior exposure to the test substance? There is a lot of uncharted ground with endocrine disrupters, and we cannot yet answer this question. An analysis by Barnthouse, Suter and Bartell (1988) showed that full chronics, i.e., those encompassing the organism's life cycle, were more sensitive than early life stage tests. However, I would not extend the chronic testing unless documented surface water or sediment exposures exceed the 600 µ/L chronic threshold that the Toxicological Review calculated (i.e., Chapter 7, Screening Ecological Risk Assessment for Perchlorate).

The second caveat concerned Nzegung's phytoremediation study. The authors of the Toxicological Review noted the lack of toxicant renewal, and suggested that bioaccumulation probably was underestimated. The lack of perchlorate renewal seems to be a significant design flaw, and it potentially may have underestimated significantly both degradation and bioaccumulation. Future tests should use a renewal or continuous flow technique to better simulate groundwater or irrigation water. The basis for biodegradation should be identified, as Nzegung proposed.

3. **Strengths and Weaknesses of Data Analyses?** The data generally were analyzed with appropriate methodologies and to the extent necessary. On important toxicological index, the acute-chronic ratio, was not measured in one set of tests, because of their design. If the BEA tests had measured the 48-hr LC50 for *Ceriodaphnia* and the 96-hr LC50 for *Pimephales*, it would have been possible to calculate acute-chronic ratios for both species. These ratios are very essential for estimating chronic toxicity from acute tests, and there are not enough of them

available for perchlorate. Although they were calculated by EA for the same species, duplicate values would have afforded a check on precision.

4. **Has the document adequately evaluated the results of all relevant studies and the biological significance of the entire database?** In general, the existing toxicological data were comprehensively presented and interpreted by Sprenger et al. (1998). Including the test results for *Hydra*, *Bufo*, lamprey and Bringman and Kuhn's (1977) *Daphnia magna* test would strengthen the weight of evidence, unless these tests are seriously flawed. The Toxicological Review mentions these tests, but did not include them in the analyses or inferences. Reasons for their exclusion were not given. I believe these tests of other species, some very distant phylogenetically from fish and cladocerans, increase the weight of evidence suggesting that perchlorate probably is not chronically toxic to aquatic life below 600 ug/L.

3. **Completeness of Technical Documentation?** The technical documentation was excellent.

4. **Are There Any Sections That Could Be Improved?** Because of my concern about ephemeral pond-wetland exposure in arid regions, I think that more testing is needed concerning the following.

Highest Priority — Documentation of perchlorate concentrations in wetlands and small streams in arid regions near areas where significant quantities of perchlorate have been disposed. Small streams and ponds, perennial and ephemeral, in arid regions are very attractive to wildlife because these habitats and water are so limited. Such areas are subject to evaporative concentration. Thus, a combination of limited rainfall, evapoconcentration and bioaccumulation in wetland plants and invertebrates could create potentially higher exposure than assumed thus far.

Highest Priority — Perchlorate's chronic toxicity potential to nesting birds dependent on wetlands (e.g., avocets, blackbirds, and stilts). Chronic tests with appropriate surrogate avian species may be necessary. Rationale: A variety of birds are very water-dependent for nesting, and avocets and stilts and blackbirds are often found using ponds throughout the West. In some locations (San Joaquin, CA; Great Salt Lake, UT), they have sometimes been placed at risks from selenium that has bioaccumulated in their invertebrate food. Use of ephemeral ponds and perennial ponds subject to extensive evapoconcentration seem to increase risk. Wildlife using evaporation ponds and wetlands draining areas with significant perchlorate in soils could be at risk, depending on water concentration, magnitude of bioaccumulation, and toxicity.

Highest Priority	Bioaccumulation potential in wetland plant species. Rationale: Bioaccumulation factors of up to 25-fold background were reported in the Toxicological Review, and Nzengung (1998) reported ClO_4 residues in leafs to be 3536 mg/kg in cottonwood, 813 mg/kg in willow, and 641 mg/kg in Eucalyptus. Residues in leaves were the highest found in plants.
Highest Priority	Perchlorate's chronic toxicity potential to a rodent. The mammalian tests of rodents conducted to support the human health risk assessment will satisfy this data requirement. Rodents will also be attracted to the wetlands described above.
Highest Priority	Degradation potential in wetland plant species (include microbial role). This study can be dovetailed to the plant bioaccumulation study. It is designed to index persistence.
Low Priority	Sediment toxicity to the freshwater amphipod, *Hyallela.* I accept the Toxicological Review's argument that water column toxicity should be indicative of sediment toxicity, so believe sediment testing constitutes a low priority. Because of the concern about exposure in evaporative ponds and small wetlands, sediment exposure of invertebrates will occur, however.

No other recommendations appear to have sufficient priority to mention. I do not believe the evidence suggests risks are sufficient to justify the following studies that were recommended by the Toxicological Review:

- Effects on aquatic plants: I saw no indication of phytotoxicity in the studies reviewed. Aquatic plants generally are protected by concentrations that protect aquatic life, because they are less sensitive (Kenaga and Moolenaar (1979).
- Effects on nondaphnid invertebrates: Daphnids generally the most sensitive group of aquatic invertebrates, and given perchlorate's low toxicity, further testing appears questionable.
- Effects on litter-feeding invertebrates: Rather than a toxicity study *per se*, examining bioaccumulation in species like corixids (backswimmers) and chironomids (midge flies) – abundant pond-dwelling species – may be more helpful. Some of the corixids and chironomids are tolerant species, and interest in them centers on what perchlorate residues they could contain, when eaten by birds.
- Testing in estuarine waters: The evidence does not suggest risks to large waterbodies. Groundwater concentrations are too low; groundwater flows are often relatively small relative to overall surface water flows, and dilution capacities frequently large in large streams and estuaries.

5. **Is the Document Useful for Characterizing Ecotoxicological Effects. If Not, Specify the Nature and Extent of Changes?** Yes, the data presented are sufficient to estimate risks, and the Toxicological Review did a good job of estimating risks, from a screening perspective.

SCREENING ECOTOXICOLOGICAL RISK ASSESSMENT (CHAPTER 7 OF TOXICOLOGICAL REVIEW)

1. **Have the Goals and Objectives of the Ecological Screening Analysis Been Adequately Described?** Yes. This document was written by scientists well versed in ecorisk assessment and in the interpretation/use of ecotox effect and exposure data.

2. **Does the Analysis Support the Summary and Conclusions Presented?**

   - Most of the conclusions are highly supportable.

   - I am reluctant to support the following conclusions/recommendations:

     - risks to earthworms: I do not believe a interspecies safety factor of 242 is warranted, and believe it should be much lower. Strengthening the scientific basis for any safety factor would be helpful. This safety factor resulted in earthworms being categorized as one of the most sensitive species tested, which is inconsistent with the scientific literature on their relative sensitivity.

     - Risks to herbivorous mammals: The endpoint selected (effect on iodide uptake by the thyroid) has not, to my knowledge, been linked directly to population-level effects. If it is demonstrated to be directly translatable to a population-level effect, then it is acceptable. However, toxicological endpoints for aquatic life and wildlife must be population-level effects. For aquatic life and wildlife, sublethal physiologic or biochemical changes (biomarkers) are not accepted as appropriate surrogates by EPA (Stephan et al. 1985) or the scientific community (Gentile and Slimak (1992). Effects on growth, survival and reproductive success are the commonly accepted population-level measurement endpoints.

3. **Are Relevant and Important Aspects of Uncertainty Addressed Sufficiently?** Yes. The uncertainty analysis was very thorough, though quite conservative.

4. **Utility of Bioassays for Characterizing Hazard (Effects Potential):** The bioassays were acceptable for characterizing hazard, subject to the caveats mentioned above concerning the potential for chronic toxicity over the life cycle of aquatic vertebrates. I was disappointed that the screening ERA only used the *Ceriodaphnia* and fathead minnow acute and chronic data. I would have preferred it would have used the other

acute aquatic tox data in applying the Tier 2 methodology, as mentioned above. No reason for excluding data on *Hydra*, for example, was given.

FURTHER ECOTOXICOLOGICAL Testing Needs for Perchlorate

1. Will The Additional Ecotoxicological Studies Currently Underway Be Sufficient To Characterize The Ecotoxicological Potential Of Perchlorate? If not, what are data needs and why, and associated experimental designs?

The following repeats recommendations made above:

Because of my concern about ephemeral pond-wetland exposure in arid regions, I think that more testing is needed concerning the following. Details are provided in the review of Sprenger et al. (1998).

Highest Priority — Documentation of perchlorate concentrations in wetlands and small streams in arid regions near areas where significant quantities of perchlorate have been disposed. Small streams and ponds, perennial and ephemeral, in arid regions are very attractive to wildlife because these habitats are so limited, as is water. Such areas are subject to evaporative concentration. Thus, a combination of limited rainfall and evapoconcentration and bioaccumulation in wetland plants could create potentially higher exposure than assumed thus far.

Highest Priority — Perchlorate's chronic toxicity potential to nesting birds dependent on wetlands (e.g., avocets, blackbirds, and stilts). Tests with appropriate surrogate avian species may be necessary. Rationale: A variety of birds are very water dependent for nesting, and avocets and stilts and blackbirds are often found using ponds throughout the West. In some locations (San Joaquin, CA; Great Salt Lake), they have sometimes been placed at risks from selenium that has bioaccumulated in their invertebrate food. Evaporation ponds and wetlands draining areas with significant perchlorate in soils could be at risk, depending on water concentration, magnitude of bioaccumulation, and toxicity.

Highest Priority — Bioaccumulation potential in wetland plant species. Rationale: Bioaccumulation factors of up to 25 x background were reported in the Toxicological Review, and Nzengung (1998) reported ClO_4 residues in leafs to be 3536 mg/kg in cottonwood, 813 mg/kg in willow, and 641 mg/kg in Eucalyptus. Residues in leaves were the highest found in plants.

Highest Priority	Perchlorate's chronic toxicity potential to a rodent. The mammalian tests of rodents conducted to support the human health risk assessment will satisfy this data requirement. Rodents will also be attracted to the wetlands described above.
Highest Priority	Degradation potential in wetland plant species (include microbial role). This study can be dovetailed to the plant bioaccumulation study. It is designed to index persistence.
Low Priority	Sediment toxicity to the freshwater amphipod, Hyallela. I accept the Toxicological Review's argument that water column toxicity should be indicative of sediment toxicity, so believe sediment testing constitutes a low priority. Because of the concern about exposure in evaporative ponds and small wetlands, sediment exposure of invertebrates will occur, however.

LITERATURE CITED

Barnthouse, L.W., G.W. Suter II, and S.M. Bartell. 1988. Quantifying risks of toxic chemicals to aquatic populations and ecosystems. Chemosphere 17(8):1487-1492.

Gentile, J.H. and M.W. Slimak. 1992. Endpoints and indicators in ecological risk assessments. pp. 1385-1397. In: [D.H. McKenzie, D.W. Hyatt, and V.J. McDonald, Editors]. Ecological Indicators. Elsevier Applied Science, New York.

Kenaga, E.E. and J. Moolenaar. 1979. Fish and Daphnia toxicity as surrogates for aquatic vascular plants and algae. Environmental Science and Technology. 13:1479-1480.

Stephan, C.E., D.I. Mount, D.J. Hansen, J.H. Gentile, G.A. Chapman, and W.A. Brungs. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. U.S. Environmental Protection Agency, Washington, D.C. NTIS No. PB85-227049. 98 pages.

Review of Genotoxicity Assays

Prepared by David Brusick, Ph.D.
Covance Laboratories, Inc.
Vienna, VA 22182

GENETIC TOXICOLOGY

Review of individual studies initiated since May 1997 and their impact on the Report

Strengths and weaknesses of the experimental designs and validity of the interpretations contained in the report: Three studies were included in the series of assays for Perchlorate (Salmonella reverse mutation, Mouse Lymphoma assay and the Mouse micronucleus). This is an acceptable battery for most chemical screening; however, this array of tests is not as comprehensive as one might want to have for this chemical considering its extensive exposure to humans.

Salmonella reverse mutation assay - the assay reported was a very basic design. Most guidelines recommending this test include additional tester strains (e.g. TA104 and/ or *E. coli* WP2 *uvr*A). Addition of these strains would have strengthened the assay and gave stronger support for assuming that perchlorate is non-genotoxic. Another method which is sometimes conducted to detect weak agents is the suspension assay which gives better contact between test material and the target organisms. The current assay meets minimum specifications.

Mouse Lymphoma - the design of the study and the conduct of the study were both inadequate. The design should have specific criteria for selecting the high dose on the basis of toxicity. Current guidelines (OECD) require that the high dose be selected at 80-90% toxicity. In this assay neither the -S9 or +S9 trials were even close to appropriate dose range bases on the preliminary toxcity study. The high dose should have been set at 5000 ug/ml for the -S9 trial and at about 2000 ug/ml for the +S9 trial. The retest of the -S9 trial should have been set at a 24 hour exposure period based on current protocols. The results from this study are not acceptable for classifying perchlorate and must be repeated. A retest of perchlorate is apparantly in progress at this time.

The mouse micronucleus test - the study design is in complete. For example, the design description does not indicate if the 1000 mg/kg dose was delivered as a single dose or was given as 2 or more doses. If the dose was given as a single acute dose (which is what I conclude), there should have been two sample times; one at 24 hours and one at 48 hours. In this study the sample time (only 1) was not stated. I am also concerned about dose selection. The logic used to go from a dose of 2000 mg/kg which was moderately toxic to the high dose in the test of 1000 mg/kg was not well supported. There was no evidence for toxicity at 1000 mg/kg which is contrary to the recommendations from OECD. In order to provide the best possible support for negative genotoxicity, I would have selected a high dose of 2000 mg/kg (or at least 1500 mg/kg) and two lower dose levels at 50% and 25% of the high dose. The other potential issue resulting

from the dose selection is that one has negative results with no evidence that the test material even reached the target cells (bone marrow). Consequently, the results from this assay are weak support for reaching a conclusion of non-genotoxic and should be repeated with a better design and higher dose levels.

Hazard Characterization

Based on my concerns for at least two of the 3 tests conducted to define the genetic toxicity of perchlorate, it is my opinion that the classification of perchlorate as non-genotoxic is not well supported. This concern my have some influence on discussions of the type of risk analysis used in future evaluations.

Further Needs for Testing

While the general structure and toxicity of perchlorate do not fit with chemicals associated with DNA activity, I believe that additional testing for DNA damage should be performed including the repeat of the Mouse Lymphoma assay and the mouse micronucleus assay. Based on completion of those repeats, additional studies may be appropriate.